

Forum for kvantitativ metode (SPS)

Sosiale interaksjonseffekter:
Hvordan kan de identifiseres?

Andreas Kotsadam

Outline

- What is a peer effect?
- How is it usually measured?
- How should we measure it?
- How do we do it in practice?
- Longer examples of peer effects from our projects in the Army.
- And/or Neighborhood effects.

Peer effects

- Very popular in the field of education: 5 most popular articles have been cited more than 2000 times.
- Broad definition: Any externality in which peer's backgrounds, behavior, or outcomes affect an outcome.

Does having smart classmates *make* you smarter?

- Possible mechanisms: Direct learning, teacher teach at higher level, less disruption in classroom, increased motivation via competition.

Types of peer effects in the classroom

Model	Homogenous effects?	Description
Linear-in-means	Yes	Only the mean of peers background or outcomes matters
Bad apple	Yes	One disruptive student harms everyone
Shining light	Yes	One excellent student provides great example for all
Invidious Comparison	No	Outcomes are harmed by the presence of better achieving peers
Boutique/tracking	No	Students perform best when surrounded by others like themselves
Focus	Yes	Classroom homogeneity is good, regardless of student <i>i</i> 's ability relative to the homogenous classmates
Rainbow	Yes	Classroom heterogeneity is good for everyone
Single crossing	No	Positive effects from high ability classmate is weakly monotonically increasing in own ability

Other examples of peer effects

- Does knowing immigrants *make* people less racist?
- Does exposure to female leaders *reduce* bias?
- Does hanging with drug users *make* you use drugs?
- Does having a daughter *affect* your political preferences?
- Do average schooling level *affect* individual earnings/fertility etc? I.e. Is there a social return to schooling?
- Is voting contagious?
- Does the share of black students *cause* the phenomenon of acting white?

Peer effects

- The main challenge is separating self selection from actual effects.
- Loosely speaking we are trying to estimate the effect of groups on individuals.
- Simply using regression analysis to do this is difficult.

Definitions from Manski (1993)

- Endogenous effects: Those that emanate from from peer's current behavior/outcomes.
- Exogenous effects: Those that emanate from from peer's backgrounds.
- Correlated effect: individuals behave similarly because they have similar characteristics or face similar institutional environment (E.g. bad teacher, poor neighborhood).

The school achievement of a youth

- Endogenous effect: If, all else equal, individual achievement *varies with average results* of classmates, co-ethnic schoolmates etc.
- Exogenous effect: if it varies with the socioeconomic composition of the group.
- Correlated effects: If it varies by a nonsocial phenomenon e.g. by teacher.

Most common model

- The linear-in-means model:

$$Y_i = \alpha + \beta_1 * \bar{Y}_{-i} + \gamma_1 * X_i + \gamma_2 * \bar{X}_{-i} + \varepsilon_i$$

Average outcome of peers
("Endogenous effects")

Average background of peers
("Exogenous effects")

What is being measured (**Assuming identification is ok!!!**)

- The model constrains the size of either peer effect (b_1 or g_2) to be the same regardless of where the student falls within the distribution of student background or ability.
- And by definition all peer effects work through the mean. Effects from any other aspects of the distribution of the peers' background are ruled out.

Potentially important critique

- Carell et al. (2011) use well identified results from the US Air Force Academy to set up the optimal group.
- "Our results suggest that using reduced form estimates to make ... policy predictions can lead to unanticipated outcomes"
- No shit: They failed miserably and got negative treatment effects.
- The problem was that within the "optimal peer groups" students self selected into sub-groups.

Non-linearity

- Hoxby (2000) and Hoxby and Weingarth (2005) argue that the linear in means model of peer effects is perhaps not the right one or even the most interesting one.
- Convincing arguments for this in educational research: The individuals own position in the ability *distribution* as well as changes in the group *distribution* is what matters.

Why not interesting?

- From a social welfare point of view the model is not that interesting since the model constrains the net effect from reassignment of peers to different classrooms or groups to be zero.

Easy to solve with an expanded model

- Duncan et al. (2005) and Sacerdote (2001) group student i into one of several possible categories and i 's peers into categories and then include in the regression all possible interactions of student i 's type and i 's roommate's type.

$$Y_i = \alpha + \lambda_1 * (D_i = 0 * D_{-i} = 0) + \lambda_2 * (D_i = 1 * D_{-i} = 1) + \lambda_3 * (D_i = 0 * D_{-i} = 1) + \varepsilon_i$$

Hoxby and Weingarth (2005)

- Divide students into deciles of past test score performance. They then interact student i 's decile of previous score with the percent of i 's peers (classmates) falling into each of the 10 deciles.
- This generates 100 interaction terms. The coefficients on these interaction terms allow the authors to test a wide variety of hypotheses about peer effects.
- Must be viewed as exploratory though.

Identification

- The fundamental challenge for the peer effects literature!
- At least three reasons why OLS is problematic:

$$Y_i = \alpha + \beta_1 * \bar{Y}_{-i} + \gamma_1 * X_i + \gamma_2 * \bar{X}_{-i} + \varepsilon_i$$

3 problems

- The reflection problem: Since student i 's outcome (Y_i) affects his peers' mean outcome (Y_{-i}) and vice versa, B_1 is subject to endogeneity bias.
- Correlated effects: peers self select into peer groups or classrooms in a manner that is unobserved to the econometrician. Bias in the estimated magnitude of peer effects B_1 and g_2 .
- Separate identification of B_1 and g_2 is difficult since peer background itself affects peer outcome. Even if one has exogenous variation in peer background characteristics (as in many of the roommates papers), that does not imply that both coefficients are separately identified.

Social multipliers

- Note that endogenous effects (those from peers' average outcome $(Y-i)$) have the potential for social multipliers since a small change for student i will affect the peer group which will then reflect back to student i , and so on.
- One way would be to combine good identification of peer effects with some random treatment within groups.

How (not) to do it practice?

Recent example

- Lee and Gaddis (2013) AJS paper.
- Social scientists and policymakers generally share the widely held belief that impoverished contexts have harmful effects on children.
- Does classroom poverty affect test scores?

The problem

- “The scholarly consensus on contextual effects, however, rests largely upon cross-sectional studies, which do not provide a strong basis for causal inference.
- Selection bias, perhaps the most important threat to the validity of point-in-time studies ...
- The only way to eliminate such correlations is to assign individuals randomly to groups, and this is impossible with observational data. “

They acknowledge solutions

- "... propensity score matching and weighting, comparison of sibling and neighbor correlations, fixed effects, instrumental variables, and natural experiments. Experimental evidence on the effect of changes in school and neighborhood context and academic achievement has emerged from the Moving to Opportunity program."

They use Panel data

- “This study uses longitudinal data to estimate the *effect* of exposure to a high poverty classroom on elementary and middle school students’ test scores.” (My emphasis).
- “To address endogenous self-selection based on fixed unobservables, we present student fixed effects estimates, which remove between-student confounding. This approach controls for time-invariant unmeasured and mismeasured aspects of student and family background that may predict both family choice of neighborhood and school and test score achievement.”

Ok, fine, but where do the variation come from then?

- "We identify the classroom poverty effect from year-to-year variation in the poverty composition of students' classrooms.
- This changes due to school mobility and due to variations in the poverty compositions of student's assigned classrooms as they progress through grade levels in the same school.
- Because classroom poverty rates vary more between schools than within schools, school movers are somewhat more likely to experience a change in classroom poverty than students who remain in the same school. "

Do you see any problems with this?

Better way

- Including student fixed effects *and* school fixed effects and identifying peer effects using cohort to cohort variation within school.
- Or include school-by-grade fixed effects and hence use classroom-to classroom level variation within a school and grade.
- The basic concept in these papers is that the student, school, or school-by-grade fixed effects remove selection effects and allow the researcher to identify peer effects from idiosyncratic variation in peer ability.

Hoxby (2000)

- Hoxby (2000b) relies on random variation in the gender and racial makeup of peers to provide estimates of peer effects.
- She uses data from students in all Texas elementary schools in grades 3–6.
- Her strategy relies on the fact that within a school and grade level, cohort level variation in gender and racial composition is an unexpected shock to peer achievement.

The two most common methods

- 1) Random assignment of individuals to groups: E.g. Students to classes, schools, or dorms.
- 2) Use *school* and individual fixed effects .
- Sacerdote (2011) reviews results from these two approaches in great detail.
- Most papers have one source of exogeneity and do not separately identify the exogenous and endogenous peer effects. (Reduced form)

College roommates

- Sacerdote (2001), Zimmerman (2003), and Stinebrickner and Stinebrickner (2006) find that roommates' background and current achievement affect own achievement.
- Foster (2006) and Lyle (2007) find no evidence that roommates' or hallmates' background affects own college GPA.
- Using data from the U.S. Air Force Academy, Carrell, et al. (2008) examine peer effects in an unusual context in which the full peer group is known and the institution forces a great deal of peer interaction. In that setting, they find large peer effects.

College roommates

- Perhaps the more interesting result from the literature on peer effects in higher education is the fact that while academic achievement is affected modestly by roommates and dormmates, the effects on more “social” outcomes are large.

College roommates

- Duncan et al. (2005) find that males who themselves binge drank in high school have a fourfold increase in their number of college binge drinking episodes (per month) when assigned a roommate who also reported binge drinking in high school.
- Boisjoly et al. (2006) find that white students assigned a black roommate report more support for affirmative action and students assigned a high income roommate less likely to support the statement that “wealthy people should pay more taxes.”

Critique

- Stinebrickner and Stinebrickner (2006) criticize the college roommates studies on *academic* peer effects.
- They have not been looking at the right place: small effects may be due to: High ability students, roommates not peers of potential influence, what qualities matter (ability not likely to change).

Other social outcomes

- As in the college literature on peer effects in social outcomes, the peer effects on drug use, criminal behavior, and teen pregnancy for younger students are estimated to be quite large. (Gaviria and Raphael (2001), Case and Katz (1991), Kling, Ludwig, and Katz (2005)).

Angrist (2014) critique

- Angrist does not think it is possible to estimate the endogenous effects as they are driven by a common variance in outcomes and he strongly cautions against using outcome-on-outcome estimations.
- He is also skeptical to studies where individuals whose background characteristics are thought to be important are also included in the sample thought to be affected by other individuals.
- He instead argues that the most compelling evidence comes from studies whereby there is a clear separation of the individuals thought to be affected and the peers thought to provide the mechanisms for the peer effects.

Examples

- This type of design is applied in Kling, Liebman, and Katz (2007) who analyze the effects of neighborhoods on individuals randomly assigned to receive housing vouchers in the Moving to Opportunity program.
- The neighborhood effects are only estimated by using characteristics of the neighbors but the neighbors themselves do not otherwise play any role and no effects on these old neighbors is estimated.

UiO • **University of Oslo**

Peer effects, gender, and ethnicity - Evidence from experiments in the Norwegian Armed Forces



Andreas Kotsadam

21 October 2015



EQUALITY • SOCIAL ORGANIZATION • PERFORMANCE

ESOP

Outline

Does Exposure to Ethnic Minorities Affect Support for Welfare Dualism?

- Introduction

- The field experiment and empirical strategy

- Results

Exposure to female colleagues breaks the glass ceiling

- The experiments and empirical strategy

- Results

Outline

Does Exposure to Ethnic Minorities Affect Support for Welfare Dualism?

Introduction

The field experiment and empirical strategy

Results

Exposure to female colleagues breaks the glass ceiling

The experiments and empirical strategy

Results

Does Personal Contact with Ethnic Minorities Affect Support for Welfare Dualism? Evidence From a Field Experiment

- ▶ Finseraas and Kotsadam

Background

- ▶ Majority-minority conflicts can influence policy preferences and outcomes
- ▶ A generous welfare state might be more difficult to sustain if the population is more ethnically heterogeneous (e.g. Alesina and Glaeser 2004)
- ▶ Diversity can reduce welfare spending through several channels, e.g. out-group hostility, cultural differences in spending priorities, more difficult to organize interest groups, ethnic politics

Motivation

- ▶ In the US, majority-minority conflicts have long been linked to White Americans' welfare state preferences (Gilens 1995)
- ▶ Starting with Alesina et al. (2001), the last decade has witnessed a massive interest in the impact of immigration on Europeans' welfare state preferences, but empirical results are all over the place

Contribution

- ▶ We study support for welfare dualism

Contribution: Conceptual

- ▶ Most of the literature has examined the relationship between immigration and broad or abstract measures of welfare state support (e.g. public sector size)
- ▶ We question whether retrenchment of welfare benefits is a likely scenario in a developed welfare state
- ▶ We suspect that a dual welfare state where one discriminates welfare rights based on for instance citizenship, might be the first-best option for voters concerned about immigration

Contribution

- ▶ We address the issue with a research design for causal inference

Contribution: Causal inference

- ▶ While most empirical studies suggest that intergroup contact reduces intergroup prejudice (see Pettigrew and Tropp 2006), the worry that most of these results are driven by selection, reverse causality, or both, looms large in this literature.
- ▶ A handful of studies use randomly assigned peers to study attitudes towards Blacks in the US (e.g. Carrell et al. 2015), but do not study welfare state preferences.
- ▶ Dahlberg et al. (2012) is the only exception. Neighborhood effects are, however, unlikely to be generalizable to effects of interpersonal contact since physical proximity does not necessarily imply personal contact.

Contribution

- ▶ We study this topic in a context where theory offers strong expectations

Contribution: Theory

- ▶ Contact theory specifies a set of conditions for when contact with minorities will make majority members more tolerant
 - ▶ Equal status, common goals, cooperation, sanctions, friendship potential. (Pettigrew 1998)
- ▶ Difficult to derive hypotheses and interpret the results if these conditions are not met
- ▶ E.g. competition between your in-group and out-groups over scarce resources, social rights and social status can cause out-group prejudice

Our study

- ▶ We conducted a field experiment in the Norwegian Armed Forces by randomizing soldiers into rooms (and hence into exposure to minorities)
- ▶ The characteristics of the military makes it a very good context for personal exposure to reduce hostility
- ▶ The experiment, hypotheses, variable operationalizations, exact model specifications and power calculations are described in a published pre-analysis plan (AEA RCT Registry)

The field experiment and empirical strategy

The field experiment

- ▶ We conducted a survey of all incoming soldiers of the August 2014-contingent of the North Brigade of the Norwegian Armed Forces
- ▶ All soldiers meet at Sessvollmoen to go through a program of medical and psychological testing before they are boarded on planes to Bardufoss at the end of the first day
- ▶ Importantly, since this is the first day of service, they do not know each other and do not know who they will share room with

The field experiment

- ▶ We constructed a randomization procedure which randomize soldiers to share rooms during the “recruit training period” (first 8 weeks of the service).
- ▶ In these rooms they perform tasks together, such as cleaning the room for inspection each morning.
- ▶ They also serve in the same platoon and normally constitute a team within the platoon.
- ▶ This period is very strict and the soldiers have to wear uniforms 24/7 and are not allowed to sleep outside of base. As the base is remotely located this implies that soldiers spend all time with each other.

The field experiment

- ▶ At the end of the recruit training period we repeated the survey

Data: Outcomes

- ▶ Immigrants should not have the same rights to social assistance as Norwegians (1=Strongly agree, 5=Strongly disagree).
- ▶ In general, immigrants have poorer work ethics than Norwegians (1=Strongly agree, 5=Strongly disagree).
- ▶ Is Norway made a worse or better place to live by people coming to live here from other countries? (1, worse to 7, better).

Recap: Problems

- ▶ If one were to test the contact hypothesis using observational data on e.g. a network of friends, it is likely that there will be a positive bias in the estimation of the peer effect.
- ▶ For illustration, we run a set of naive regressions of the share of non-Norwegian friends in high school as well as regressions using the share of immigrants in the home municipality on our outcomes of interest.

Table: Naive regressions

	(1)	(2)	(3)
	Same rights t2	Work ethics t2	Better country t2
Panel A: Minority friends			
Minority friends	0.138*	0.156**	0.230**
	(0.074)	(0.063)	(0.109)
Observations	533	534	533
Platoon FE	Yes	Yes	Yes
Panel B: Share of immigrants in the municipality			
Share of immigrants	1.592***	0.770*	1.011**
	(0.462)	(0.408)	(0.493)
Observations	584	585	584
Platoon FE	Yes	Yes	Yes

Data: Treatment and control group

- ▶ *TREATED* equals 1 if the soldier shares room with a soldier with a non-western background (NWB)(treatment group), and equals 0 if not (control group)
- ▶ 5 percent of the soldiers have a NWB, 21 percent of the sample are treated
- ▶ We only use information on *assigned* room and, following Angrist (2014) ([▶ Details](#)) we only include WB people in the regressions.

Empirical specification

$$Y_{irt2} = \alpha_J + \beta_1 \text{Treated}_r + \beta_2 Y_{irt1} + \beta_n X_{irt1} + \epsilon_{ir}$$

where α_J is platoon fixed effects, Y_{t1} is Y at baseline (day 1), X_{irt1} refers to the vector of potential controls. SE are clustered at room.

Results

Table: Regressions of treatment status on pre-determined variables.

	Coeff	t	Standardized coeff	N
Same rights t1	-.13	1.20	-.05	589
Work ethics t1	-.16	1.50	-.07	552
Better country t1	.05	0.33	.01	552
Mother has high education	-.02	0.38	-.02	550
Father has high education	.00	0.07	.00	550
Mother is employed	-.09**	2.05	-.12	549
Father is employed	-.02	0.40	-.06	549
Parents are divorced	.00	0.01	.00	549
Plan to take higher education	.01	0.16	.01	551
IQ	-.01	0.09	-.00	601
F-test of joint significance	1.07 (p=.38)			

Note: Each row presents the results from one regression. Platoon fixed effects are included in all regressions.

t-values adjusted for room clustering. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table: Main results

	Same rights t2	Work ethics t2	Better country t2
No controls			
Treated	0.037 (0.085)	0.196** (0.085)	0.083 (0.124)
Same rights t1	0.610*** (0.039)		
Work ethics t1		0.582*** (0.046)	
Better country t1			0.635*** (0.043)
Platoon FE	Yes	Yes	Yes
Observations	534	535	534

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table: Main results

	Same rights t2	Work ethics t2	Better country t2
Control for difference in mother's employment			
Treated	0.012	0.187**	0.080
	(0.084)	(0.085)	(0.124)
Mother is employed	-0.068	-0.007	-0.152
	(0.111)	(0.116)	(0.153)
Baseline outcome	Yes	Yes	Yes
Platoon FE	Yes	Yes	Yes
Observations	531	532	531

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table: Main results

	Same rights t2	Work ethics t2	Better country t2
Full set of individual level controls			
Treated	0.000	0.187**	0.058
	(0.084)	(0.085)	(0.126)
Baseline outcome	Yes	Yes	Yes
Platoon FE	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes
Observations	522	523	522

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Robustness checks

- ▶ Ordered probit and LPM with dichotomized dependent variables ▶ Tests
- ▶ Share of minority soldiers in the room ▶ Tests
- ▶ Control for share with highly educated parents in the room ▶ Tests
- ▶ Placebo ▶ Tests
- ▶ Non-random attrition ▶ Tests
- ▶ Adjustment for multiple testing ▶ Tests, treatment heterogeneity ▶ IQ, and exploratory analysis ▶ Analysis

Conclusion

- ▶ We find quite large and statistically significant effects of personal contact on views on immigrants' work ethic.
- ▶ Contrary to our expectation, the improved view on immigrants' work ethic is not reflected in reduced support for welfare dualism.
- ▶ The same is true for views on whether immigration makes the country a better place to live.

External validity

- ▶ Although the context of our study is in part a necessity for deriving clear theoretical expectations and while it assures a strong internal validity, it restricts external validity to contexts with some similarity to ours.
- ▶ The structure of contact at workplaces, in classrooms, and in team sports are weaker and less streamlined which might imply that treatment effects from direct contact might be weaker than what we find.

Outline

Does Exposure to Ethnic Minorities Affect Support for Welfare Dualism?

Introduction

The field experiment and empirical strategy

Results

Exposure to female colleagues breaks the glass ceiling

The experiments and empirical strategy

Results

Exposure to female colleagues breaks the glass ceiling - Evidence from a combined vignette and field experiment

- ▶ Finseraas, Johnsen, **Kotsadam**, and Torsvik

Introduction

- ▶ “Fewer Women Run Big Companies Than Men Named John” (NYT March 2)

Introduction

- ▶ “Fewer Women Run Big Companies Than Men Named John” (NYT March 2)
- ▶ This vertical segregation is commonly referred to as the glass ceiling and it is blatant also in Norway:

Introduction

- ▶ “Fewer Women Run Big Companies Than Men Named John” (NYT March 2)
- ▶ This vertical segregation is commonly referred to as the glass ceiling and it is blatant also in Norway:
 - ▶ The gender gap in wages is 50 percent higher among college graduates than among full time working men and women in general, and before quotas were introduced in corporate boards only 5 percent of board members were women (Bertrand et al. 2014).

Identifying discrimination

- ▶ Such differences are likely partly due to supply side factors (preferences, hh-work, competitiveness).
- ▶ The differences may also stem from demand side discrimination.
- ▶ Identifying discrimination is difficult to do with observational data as many of the factors that may influence the valuation of a candidate are not observed by the researcher.
- ▶ We use a randomized vignette experiment.

Identifying peer effects

- ▶ Finding what determines discrimination is important and we have reasons to believe that exposure may reduce it:
 - ▶ Random exposure to female village leaders in India (Beaman et al. 2009) and of black and white roommates in college (e.g. Boisjoly et al. 2006) or in the US Air Force (e.g. Carrell et al. 2015) has been shown to reduce bias.
- ▶ Whether peer exposure to women reduces the amount of discrimination has not been tested before.
- ▶ Challenging to test peer effects due to homophily.
- ▶ We conduct a field experiment where we randomize exposure to female colleagues.

The paper in a nutshell

- ▶ Setting: Norwegian Armed Forces. Conscripts during the first 8 weeks of service (Boot camp).
- ▶ Vignette experiment: Evaluate fictive male/ female squad leader candidates.
- ▶ Finding 1: Female candidates valued less than male candidates.
- ▶ Field experiment: Female recruits randomly assigned to male rooms.
- ▶ Finding 2: Males from mixed rooms do not discriminate.

Main contributions

- ▶ Previous literature has identified a clear pattern, whereby gender discrimination covaries positively with the gender composition of the sector of employment.
- ▶ The Norwegian Armed Forces have fewer women in top positions than any other Norwegian sector, including the church (Teigen 2014).
- ▶ Our results are of interest in order to understand the advancement of women in a hyper male setting.
- ▶ We move beyond merely identifying discrimination to show that exposure reduces it.

Appendix

The field experiment

- ▶ Same setup as in the the immigrant case.
- ▶ But here we focus on the Second Battallion of the North Brigade
- ▶ For whom we conducted a Vignette experiment in order to measure discrimination.

Our Vignette experiment

- ▶ Evaluate a fictional candidate *on a scale 1-6*.
 - ▶ Advantage of using a scale
- ▶ Experimentally manipulate gender and information.
- ▶ 4 treatments randomly allocated to 413 soldiers: Ida basic, Martin basic, Ida more info, Martin more info.
- ▶ Ida/ Martin most common names for 1994-cohort.
- ▶ Ran the experiment 26th September, 2014.
- ▶ 8 sessions.

SQUAD LEADER:

The unit is choosing a new squad leader. The squad leader is the link between officers and soldiers. For some, this position can be very physically and mentally demanding. The position requires high skills. As squad leader one is responsible not just for oneself, but also for the team.

A potential candidate: Name: [Ida Johansen/ Martin Hansen](#)

- Grades from high school: 4.1 (average).
- Career plans: Does not wish to continue in the armed forces, plans to pursue higher education (civilian) in the field of economics and administration.
- Family background: Has a sister, dad is an engineer and mother is a teacher. Comes from a middle-sized city in the eastern part of Norway.
- Motivation: Thinks that serving in the armed forces is both meaningful and important.
- Physical capacity: [Among the top 20 percent in his/ her cohort \(armed forces\). Exercise regularly.](#)
- Leadership experience: [Was the leader of a youth organization.](#)

LAGFØRER

Troppen skal velge nye lagførere. Lagfører er bindeleddet mellom befalet og soldatene. For noen vil denne stillingen være meget fysisk og psykisk krevende. Det stilles høye krav for å bli lagfører. Som lagfører har man ikke bare ansvaret for seg selv, men også et lag.

En mulig kandidat

Navn: **Ida Johansen/ Martin Hansen**

- *Fysisk kapasitet:* Var blant de 20% beste på sesjonskullet. Trener regelmessig.
- *Karakter fra Videregående Skole:* 4,1 i snittkarakter.
- *Ledererfaring:* Har vært leder i en ungdomsforening.
- *Karriereplaner:* Ønsker ikke å fortsette i forsvaret, sikter mot å ta høyere utdanning i det sivile innenfor økonomi og administrasjon.
- *Familieforhold:* En søster, far er ingeniør og mor er lærer. Kommer fra en mellomstor by på Østlandet.
- *Motivasjon:* Synes tjenesten i forsvaret er meningsfull og viktig.

Ida Johansen/ Martin Hansen vil gjerne bli lagfører, marker hvor godt synes du **han/ hun** passer til den jobben: (1= svært dårlig, 6= svært godt) – sett en sirkel rundt ditt valg:

1 2 3 4 5 6

Theory and testable hypotheses

- ▶ Taste-based discrimination: Personal prejudice of agents who dislike associating with individuals of a given gender.
 - ▶ Statistical discrimination: Employers use gender to extrapolate a signal of unobserved components of productivity.
1. Discrimination if Martin is perceived as a better candidate than Ida.
 2. Statistical discrimination if more information reduces discrimination.

Evaluation of candidate

	Less info		More info	
	Ida	Martin	Ida	Martin
Mean score candidate	3.771	4.145	4.376	4.720
Standard deviation	(1.004)	(0.926)	(0.893)	(0.817)

(1=very bad, 6=very good)

No difference in background characteristics

Evaluation of candidate

VARIABLES	Info Pooled	More/ less info
Female candidate	-0.326*** (0.108)	-0.275* (0.140)
Information added		0.551*** (0.134)
Female*Information		-0.109 (0.166)
Mean of dependent variable	4.281	4.281
Observations	367	367
R-squared	0.128	0.190
Troop FE	Yes	Yes
Session FE	Yes	Yes

Notes: Standard errors clustered at the room level in parantheses.

Exposure and bias

- ▶ The discrimination literature often acknowledges that exposure is important.
- ▶ The empirical tests of this are often problematic, however.
- ▶ Correspondence analyses are sometimes combined with data on attitudes or criminal behavior in different areas (e.g. Doleac and Stein 2013) or ethnic mix of the area (e.g. Ewens et al. 2014).
- ▶ Such analyses are also problematic at a conceptual level.

Potential mechanisms (1)

- ▶ As in the immigrant example, the conditions for contact theory are ideal. In addition:
- ▶ As people tend to favor leaders that are similar to themselves, a self-fulfilling process of homosocial reproduction may occur (Kanter 1977).
- ▶ Qualitative evidence that mixed rooms reduces gender essentialist notions and increases feelings of sameness among the soldiers (Hellum, 2015).
- ▶ Hence, it is possible that intense exposure makes male soldiers perceive themselves as more similar to female soldiers and therefore less skeptical to having them as leaders.

Potential mechanisms (2)

- ▶ Another mechanism that may potentially reduce discrimination is reduced tokenism as under-representation of women in the group may lead to them being viewed as symbols or tokens.
- ▶ Previous research suggests a critical mass, whereby the perspective of the minority members and the nature of the relations in the group change qualitatively as the minority grows from a few token individuals into a considerable minority (Kanter 1977; Dahlerup 1998).
- ▶ Testable implication: Non-linear effects.

Treatment and control groups

- ▶ *TREATED* equals 1 if the soldier shares room with a female soldier (treatment group), and equals 0 if not (control group)
- ▶ 89 rooms with between 4 and 8 persons and 0-4 women
- ▶ 8 percent of the soldiers are women, 21 percent of the men are treated (share: 0-0.67, mean .07, sd 0.15). [▶ Distribution](#)
- ▶ We only use information on *assigned* room and we only include men in the regressions.

Empirical specification

$$Score_{irt2} = \alpha_J + \gamma_S + \beta_1 Room\ Treatment_r + \beta_n X_{irt1} + \epsilon_{ir}$$

where α_J is platoon (“tropp”) fixed effects, γ_S are session f.e., X refers to a vector of potential baseline controls. SE are clustered at room.

Testable hypothesis

1. Discrimination if Martin is perceived as a better candidate than Ida.
2. Statistical discrimination if more information reduces discrimination.
3. Exposure matters for discrimination if males from mixed rooms evaluate the candidate differently from males from strict male rooms.

Table: Regressions of treatment status on pre-determined variables.

	Coeff	t
Mother has high education	0.020	0.489
Father has high education	0.003	0.081
Mother is employed	0.023	0.517
Father is employed	-0.039	-0.476
Parents are divorced	0.017	0.319
Plan to take higher education	0.005	0.138
IQ	0.007	0.544
F-test of joint significance	0.03 (p=.86)	

Note: Each row presents the results from one regression. Platoon fixed effects are included in all regressions.

t-values adjusted for room clustering. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Evaluation of candidate: Peer effects.

VARIABLES	Evaluation
Female candidate	-0.430*** (0.124)
Treated	-0.230 (0.145)
Treated*Female candidate	0.513** (0.204)
Mean of dependent variable	4.281
Observations	367
R-squared	0.139
Troop FE	Yes
Session FE	Yes

Notes: Standard errors clustered at the room level in parantheses.

Non-linearities

- ▶ Regressing Score on share of exposure, the latter is highly statistically and economically significant.
- ▶ We also see that there is a clear non-linear pattern whereby having the lowest share of exposure, with only 17 percent women in the room, actually has a negative effect on the discrimination of the female candidate.
- ▶ Having at least 20 percent women in the room, however, always leads to a decline in the discrimination of the female candidate. [▶ More results](#)

Conclusions

- ▶ There are discriminatory attitudes towards women in the Norwegian army.
- ▶ The discrimination does not seem to be related to stereotypes of strength and leadership experience.
- ▶ Living together with female recruits makes the discrimination disappear.

External validity

- ▶ Military service is mandatory for men in Norway, but conscription is based on need, and only about one in six men are needed in duty.
- ▶ Since 2010, screening and testing for military service has been mandatory for both sexes, but women serve on a voluntary basis.
- ▶ Hence, both the men and the women are selected based on ability and motivation, and the women more so.
- ▶ This is probably a fact in all male dominated settings, however.

Other military projects

- ▶ IAT
- ▶ Educational aspirations
- ▶ Voting and political attitudes
- ▶ Intermixing institutions
- ▶ Teamwork
- ▶ Games: Trust, cooperation, competition, and risk.

Well known problems

- ▶ Correlated effects
- ▶ The reflection problem
- ▶ Separate identification is difficult since peer background itself affects peer outcome

[▶ Back](#)

More critique

- ▶ Angrist (2014) strongly cautions against using outcome on outcome estimations as they are driven by a common variance in outcomes
- ▶ He is also skeptical to studies where individuals whose background characteristics are thought to be important are also included in the sample thought to be affected by other individuals
- ▶ He instead argues that the most compelling evidence comes from studies whereby there is a clear separation of the individuals thought to be affected and the peers thought to provide the mechanisms for the peer effects

[▶ Back](#)

Robustness tests

- ▶ Ordered probit and LPM with dichotomized dependent variables
- ▶ Share of minority soldiers in the room
- ▶ Control for share with highly educated parents in the room
- ▶ Placebo
- ▶ Non-random attrition

▶ Back

Table: Robustness checks

	Work ethics t2	Better country t2
Panel A: Ordered probit regressions		
Treated	0.274**	0.093
	(0.119)	(0.117)
Platoon FE	Yes	Yes
Observations	535	534

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ [▶ Back](#)

Table: Robustness checks

	Same rights t2	Work ethics t2	Better country t2
Panel B: Linear probability models of binary dependent variables			
Treated	-0.012 (0.041)	0.077* (0.042)	0.093** (0.046)
Platoon FE	Yes	Yes	Yes
Observations	534	535	534

[▶ Back](#)

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

same rights and *work ethics* are recoded to binary indicators of support for by collapsing the categories “disagree” and “disagree strongly”, while better country is dicotomized by recoding categories 5-7 to 1 and the others to 0. ***

$p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table: Robustness checks

	Same rights t2	Work ethics t2	Better country t2
Panel C: Share of minority soldiers in the room			
Treated Share	0.213	0.713*	0.205
	(0.391)	(0.384)	(0.474)
Platoon FE	Yes	Yes	Yes
Observations	534	535	534

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

▶ Back

Table: Robustness checks

	Same rights t2	Work ethics t2	Better country t2
Panel A: Control for share of high educated fathers			
Treated	0.036 (0.085)	0.204** (0.084)	0.078 (0.123)
Share of high educated fathers	-0.055 (0.179)	0.270 (0.172)	-0.197 (0.266)
Platoon FE	Yes	Yes	Yes
Observations	534	535	534

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ [▶ Back](#)

Table: Robustness checks

	Same rights t2	Work ethics t2	Better country t2
Panel B: Share of high educated fathers w/o treated			
Share of high educated fathers	-0.059	0.250	-0.204
	(0.179)	(0.174)	(0.263)
Platoon FE	Yes	Yes	Yes
Observations	534	535	534

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ [▶ Back](#)

Table: Tests for non random attrition

	(1)	(2)	(3)	(4)	(5)
	Treatment	Treatment	Outcome	Outcome	Outcome
Treated	0.014 (0.048)	0.018 (0.051)			
Better country t1			-0.013 (0.011)		
Work ethics t1				-0.007 (0.015)	
Same rights t1					-0.012 (0.013)
Observations	899	766	828	828	828
Platoon FE	Yes	Yes	Yes	Yes	Yes
Individual controls	No	Yes	No	No	No

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant. *** $p < 0.01$, **

$p < 0.05$, * $p < 0.1$

Table: Placebo regressions

	Equality	Equality
	not imp	not imp
	dummy	dummy
	Men	All
Treated	-0.072 (0.105)	0.008 (0.026)
Baseline Y	Yes	Yes
Platoon FE	Yes	Yes
Observations	349	535

▶ Back

Note: Robust standard errors adjusted for clustering on room. All regressions include a constant. *** $p < 0.01$, **

$p < 0.05$, * $p < 0.1$

Adjusting for multiple outcomes

- ▶ P-values are misleading when testing several hypotheses on the same data: The more hypotheses you test, the more likely it is that you find a statistical significant difference by chance
- ▶ In our case, the Benjamini and Hochberg (1995) adjusted 10 percent significance level for work ethic is then $.10/4 = .025$
- ▶ The treatment effect on work ethic is borderline significant at the ten percent level [▶ Back](#)

Table: Treatment heterogeneity

	(1)	(2)	(3)
	Same rights t2	Work ethics t2	Better country t2
Treated high ability	0.242 (0.152)	0.129 (0.146)	-0.063 (0.234)
Treated low ability	0.016 (0.096)	0.181* (0.100)	0.065 (0.156)
F-test of diff high-low	1.8 (p=.19)	0.1 (p=.75)	0.25 (p=.62)
Baseline outcome	Yes	Yes	Yes
Platoon FE	Yes	Yes	Yes
Observations	436	437	436

Note: Robust standard errors adjusted for clustering on room in parentheses.

All regressions include a constant. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Exploratory analysis

- ▶ Another type of heterogeneity is to investigate the effects for different types of individuals
- ▶ We conduct exploratory analyses that were not part of the pre-analysis plan.
- ▶ These findings should therefore be interpreted as suggestive or as hypotheses to be tested in the future. [▶ Back](#)

Exploratory analysis

- ▶ We start by investigating the effects for different types of people based on their prior exposure to immigrants.
- ▶ We interact the treatment with a variable for the share of immigrant friends during last year in high school.
- ▶ And with a dummy for municipalities with above median shares of immigrants in the population.
- ▶ Unclear what to expect as there are many different potential mechanisms (see e.g. Wessel 2009). [▶ Back](#)

Findings

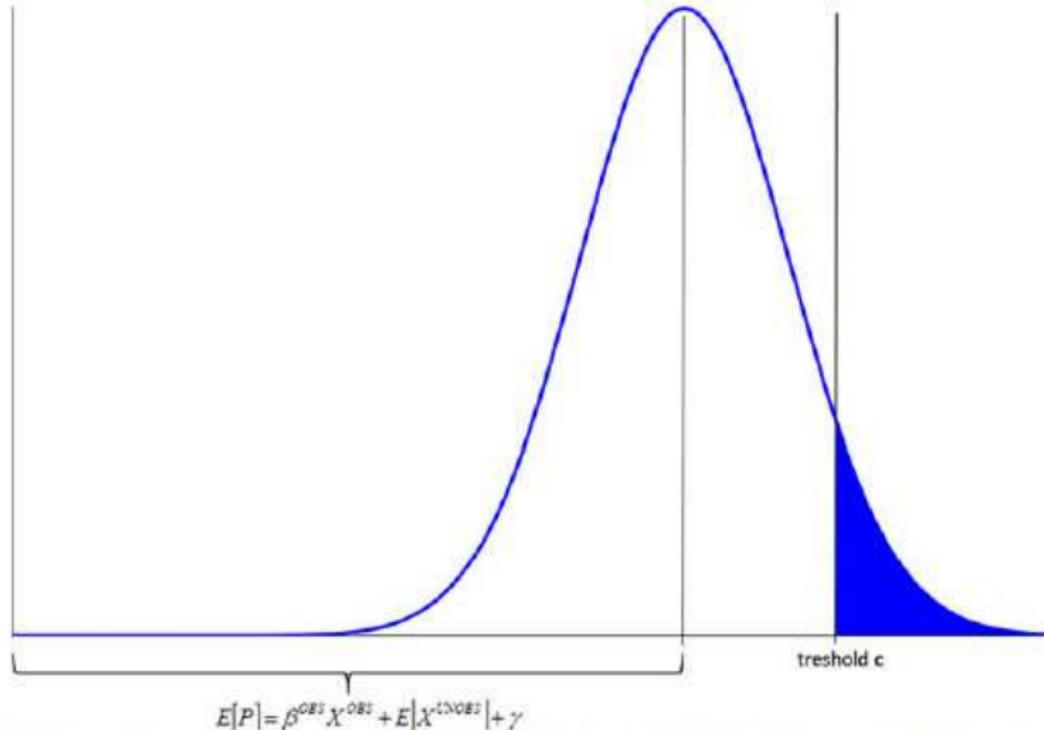
- ▶ The treatment effect is larger for individuals coming from municipalities with a higher share of immigrants.
- ▶ This is consistent with a view that the individuals from such municipalities are updating their previous misconceptions regarding immigrants work ethics that they had gained before in their (perhaps segregated) exposure to immigrants.
- ▶ Consistency is not proof, however, and there are many other differences between the two groups of municipalities.
- ▶ No differential effect for share of immigrant friends.

▶ Back

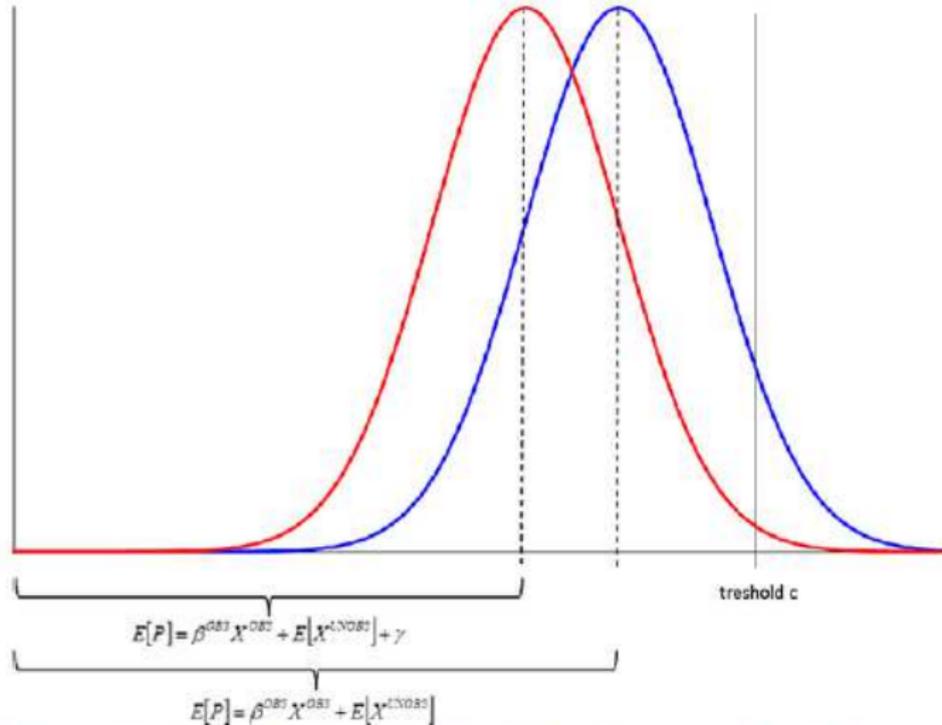
HS critique

- ▶ Differences in the distribution of characteristics can generate differences in non-linear outcomes (Heckman and Siegelman 1993).
- ▶ This is problematic, since the magnitude of the discrimination depends on the level of standardization of the job applications.
 - ▶ If candidates are matched on characteristics at an average level that is low relative to the threshold for hiring, the more heterogeneous group will have a higher share that exceeds the threshold.

Carlsson et al. 2014

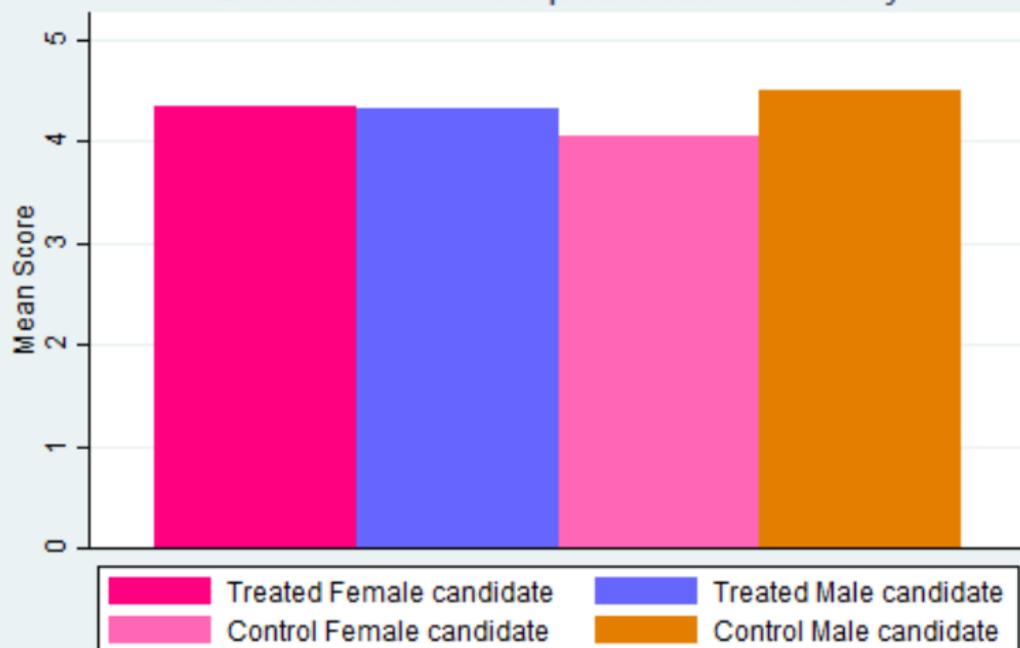


Notes: The vertical axis measures density and the horizontal axis measures productivity. The shaded area is the probability of being invited to a job interview (i.e., the probability of passing the threshold c).



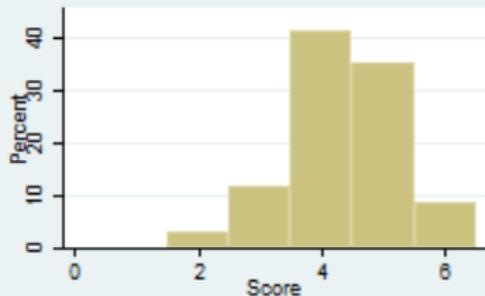
Notes: The density curve for the discriminated group is shifted to the left, either as a result of preference-based discrimination or statistical discrimination due to a perceived group difference in the mean of unobserved variables. The discriminated group is less likely to pass the threshold (compare the areas to the right of the threshold under the density curves).

Evaluations of the squad leader. Men only

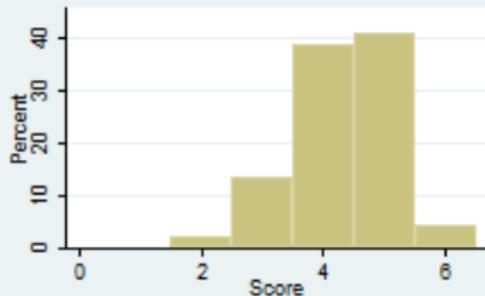


Only men included in the sample. Treated refers to sharing room with women

Treated Male with Female candidate



Treated Male with Male candidate



Control Male with Female candidate



Control Male with Male candidate

[▶ Back](#)

Table: Share of women in the room for treated soldiers. [▶ Back](#)

Share of women in room	Number of exposed men	Percent
17 % women in the room	5	6.41
20 % women in the room	4	5.13
25 % women in the room	11	14.10
29 % women in the room	9	10.26
33 % women in the room	35	44.87
50 % women in the room	13	16.67
67 % women in the room	2	2.56
Total	78	100

More Results

- ▶ Gender of respondent
- ▶ Statistical vs taste based with gender of respondent

▶ Back

VARIABLES	(1)	(2)
	Baseline	No information difference
Female candidate	-0.285*** (0.104)	-0.325*** (0.107)
Female respondent		-0.107 (0.293)
Female*Female candidate		0.641* (0.346)
Dep	4.306	4.306
Observations	398	398
R-squared	0.125	0.134
Troop FE	Yes	Yes
Session FE	Yes	Yes

Notes: Standard errors clustered at the room level in parantheses.

VARIABLES	(1)	(2)
	Treatment	Information and Treatment
Fem	-0.430*** (0.124)	-0.350** (0.164)
Info		0.538*** (0.133)
Fem*Infor		-0.130 (0.174)
Treated	-0.230 (0.145)	-0.163 (0.140)
T*fem	0.513** (0.204)	0.359 (0.223)
T*Fem*Info		0.141 (0.278)
R-squared	0.139	0.198
Troop and session FE	Yes	Yes

Notes: Standard errors clustered at the room level in parantheses.