

Exploring the sources of the gender pay gap among academics in Sweden and Norway.¹

Karin Halldén² and Kjersti Misje Østbakken³

Introduction

There has been a steady increase in the number of doctoral degrees awarded, especially among women over the last decades.⁴ The female share among PhDs have increased steadily and in we have reached gender balance⁵ among employed PhDs in both Sweden and Norway – both inside and outside academia. As shown in table 1, the female share among PhD educated individuals is 0.44 in Sweden and 0.43 in Norway. The female share is identical inside and outside academia in Sweden, while the female share is 0.44 inside academia in Norway and 0.42 outside academia. Still, research on the gender pay gap among academics is scarce, and very little is known about the gender pay gap among PhDs outside academia. As most PhDs in both countries are employed *outside* academia, it is vital to investigate the gender pay gap and explore the sources of the pay gap in both sectors.

Table 1: Shares of PhD educated men and women in Sweden and Norway (2017).

	Sweden				Norway			
	Women	Men	Observations	Share	Women	Men	Observations	Share
All	0.44	0.56	44 403		0.43	0.57	26 983	
Inside academia	0.44	0.56	16 866	0.38	0.44	0.56	9189	0.34
Outside academia	0.44	0.56	27 537	0.52	0.42	0.58	17 794	0.56

¹ This report documents the empirical analyses that are provided in the fact sheet: *The Gender Pay Gap Among PhD Holders in Sweden and Norway*

² Swedish Institute for Social Research/Mälardalens universitet: karin.hallden@sofi.su.se

³ Institute for Social Research: k.m.ostbakken@samfunnsforskning.no

⁴ This research is a part of “Nordic Centre for Research on Gender Equality in Research and Innovation” (NORDICORE), financed by NordForsk.

⁵ We define gender balance as a female/male share of between 0.4-0.6.

Utilizing register data from 2017, we study the gender pay gap among PhD holders in Sweden and Norway – countries internationally known for high levels of gender equality. We also analyse to what extent a potential wage gap could be attributed to observable factors, such as age/experience, family status and industrial- and occupational segregation, and to what extent there is a parenthood wage penalty. Our analysis contributes with knowledge on potential diverging roles of family obligations for male and female academics, inside and outside academia. Comparing male and female PhDs is a particularly compelling case since they have invested vastly in human capital, are highly work committed and specialized.

Data and variables.

The empirical analyses are based on Swedish and Norwegian administrative register data on PhDs educated individuals in 2017. For Norway the data includes administrative data organized by Statistics Norway such as linked employer-employee register (*A-ordningen*), tax records, family- and educational registers. The Swedish data comprises the university/university college register (*Personal vid universitet och högskolor*) with added registry data from LISA (*Longitudinell integrationsdatabas för sjukförsäkrings- och arbetsmarknadsstudier*).

The year of observation is 2017. Our sample consists of PhD educated individuals aged 25-65, who are registered as residents January 1st and in registered employment in November. Since academic employees often hold several part-time positions, we only include jobs that amount to at least 50 percent or more of a full-time job. Our data include information on income, field of study, whether they were employed inside or outside academia as well as family circumstances and specific individual characteristics.

We define academia as universities and university colleges only. This is a somewhat conservative definition for Norway, since external research institutes constitute a sizeable share of academic researchers and employees with PhDs. The chosen definition is a compromise to achieve comparability with Sweden. As a robustness check, we provide results for Norway with research institutes as part of the academic sector in table A1 in the Appendix. The results are qualitatively similar, and our main conclusion does not change.

In the regression analyses, we have included a set of variables that are available in both datasets and categorized them to ensure comparability across countries.

Table 2: Included variables and definitions. Sweden and Norway

Variable	Definition
Gender	Dummy variable: <i>Male (ref); Female</i>
Earnings	<i>Yearly taxable earnings</i> . This includes all earnings from all jobs (if individual has several employment spells within a year).
Native/Swedish resident	Dummy variable: <i>Foreign born/foreign resident (ref); Individuals born in Norway and/or by a mother residing in Norway / Swedish residents</i> .
Married	Dummy variable: <i>Non-married (ref); Married or registered partner</i>
Children	Four dummy variables: <i>No children (ref); One child; Two children; Three or more children</i> .
Academia	Coded into two dummy variables: <i>Universities and university colleges; all other industries</i> ¹ (ref).
Years since graduation	Coded into five dummy variables: <i>0-4 (ref); 5-9; 10-14; 15-19; 20 +</i>
Field of study	Coded into six dummy variables: <i>Education (ref)</i> (including “Teacher training”, “Education science”); <i>Humanities and social sciences</i> (including “Humanities and Arts”, “Arts”, “Humanities”, “Social and behavioural science”, “Social services” and “Journalism and information”); <i>Business, economics and law</i> (including “Business and administration” and “Law”); <i>Science</i> (including “Life sciences”, “Physical sciences”, “Health”, “Veterinary”, “Mathematics and statistics”, “Computing”, and “Environmental protection”); <i>Engineering, manufacturing and construction</i> (including “Engineering and engineering trades”, “Manufacturing and processing”, “Architecture and building” and “Agriculture, forestry and fishery”); and <i>Other</i> (including “Personal services”, “Transport services”, “Security services” and “Unknown/Not specified”)
Position (inside academia)	Coded into three dummy variables: <i>Administrative position; first position (post-doc/associate professor/researcher etc); Professor</i>
Occupation	Coded into 10 dummy variables: <i>Legislators, senior officials and managers; Professionals (ref); Technicians and associate professionals; Clerks; Service workers and shop and market sales workers; Agricultural, forestry and fishery workers; Craft and related trades workers; Plant and machine operators and assemblers; Elementary occupations; Armed forces and unspecified</i> .
Industry	Coded into 15 dummy variables: <i>Agriculture, forestry and fishing; Mining and quarrying, Manufacturing; Electricity, gas, steam and air condition supply, Water supply, waste management and remediation activities; Construction; Wholesale and retail trade, repair of motor vehicles; Transport and storage; Accommodation and food services activities; Information and communication; Financial and insurance activities; Real estate activities, Professional, scientific and technical activities, Administrative and support service activities (ref); Public administration and defence; Education; Human health and social work activities; Arts, entertainment and recreation; Other activities</i> .
Full-time	Dummy variable: <i>Part-time (less than 30 hours per week, ref); Full-time (30 hours or more per week)</i> *Available in Sweden inside academia only, available in Norway for the full sample.

Notes: 1) Including research institutes in Norway.

Empirical strategy

Our analytical approach is to first estimate the gender pay gap among PhD's estimating equation (1) using OLS.

$$(1) \quad \ln y_i = \alpha_0 + \beta_1 \text{Female} + \beta_2 x_i + \varepsilon_i$$

where $\ln y_i$ is the log of yearly individual earnings and x_i is a vector of observable characteristics. The coefficient of interest is β_1 , which is our measure of the gender pay gap. We report the point estimates in table 2.⁶

First, we estimate the unadjusted gender pay gap by including a female dummy only in equation (1). Second, we estimate the so-called adjusted pay gap by including observable characteristics. The analysis is done for all PhD's and split into PhD's working in academia and PhD's working outside academia.

The variables and model specifications are chosen to ensure comparability between Sweden and Norway, and we have also aimed for comparability between the analysis inside and outside academia. Due to structural differences between the two sectors in both countries and country differences in data availability, there are minor differences in model specifications inside and outside academia, and we therefore advice readers of this note to bear them in mind in comparisons of the gender pay gap inside academia and outside academia.

Subsequently, we decompose the gender pay gap in a counterfactual manner using the Oaxaca-Blinder decompositions (Blinder, 1973; Oaxaca, 1973). The procedure divides the gender pay gap into two parts: one that is "explained" by group differences in personal characteristics, and one that is "unexplained", i.e., that is unaccounted for by these characteristics. The unexplained part has historically been used as a measure of discrimination, but it also subsumes the effects of unobserved differences between men and women on the gender pay gap. The decomposition is performed using the STATA-command Oaxaca (Jann, 2008). We use a twofold decomposition where we assume that wage discrimination is directed at women only. The analysis is done for all PhD's and split into PhD's working in academia and PhD's working outside academia.

Last, we investigate the parenthood penalty/premium in equation (2) by adding an interaction term between the dummy variable for children and female, and then calculate the pay gap between: a) mothers and comparable women without children and b) fathers and comparable men without children. The analysis is done for three groups based on the age of the youngest child: Children aged 0-5, 6-12 and 13-17.

$$(2) \quad \ln y_i = \alpha_0 + \beta_1 \text{Female} + \beta_2 \text{Child}_t + \beta_3 \text{Female} * \text{Child}_t + \beta_4 x_i + \varepsilon_i$$

⁶ Note that we transform the point estimate to percentages by the following formula: $(e^{\beta_1} - 1) * 100$ in the fact sheet from this study.

Results

The estimated coefficients from regressions on equation (1) are presented in table 3. We report the point estimates and the corresponding standard errors. The unadjusted gender pay gap corresponds to what we often call the observed gap. This pay gap is unconditional on observable individual characteristics that may affect the wage level, such as age. In the adjusted pay gap, we have controlled for gender differences in observable characteristics. These pay gaps are often referred to as a gender pay gap between relatively similar men and women. We also provide point estimates where we include a dummy variable for working full time. This variable is only available for academia in Sweden and is therefore left out of the main specifications.

Table 3: The gender pay gap: Point estimates from separate regressions. Sweden and Norway (2017).

	Sweden			Norway		
	All	Inside academia	Outside academia	All	Inside academia	Outside academia
Unadjusted	-0.167*** (0.0007)	-0.119*** (0.006)	-0.196*** (0.010)	-0.223*** (0.006)	-0.102*** (0.008)	-0.285*** (0.005)
Adjusted	-0.122*** (0.006)	-0.057*** (0.004)	-0.143*** (0.009)	-0.151*** (0.005)	-0.059*** (0.007)	-0.183*** (0.007)
Adjusted incl fulltime		-0.049*** (0.004)		-0.137*** (0.005)	-0.060*** (0.007)	-0.167*** (0.007)
Observations	44 403	16 866	27 537	26 983	9189	17 794

Note: Standard errors in parenthesis. The three models include the following sets of control variables; *All*: female, age, age², born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), industry (dummy variable). *Inside academia*: female, age, age², born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), position (dummy variable). *Outside academia*: female, age, age², born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), industry (dummy variable), occupation (dummy variable). * p<0.1, ** p<0.05, *** p<0.01.

The first panel of table 3 show that there is a significant observable gender pay gap among PhDs in both Sweden and Norway. The overall gap is larger in Norway, mainly driven by a larger gender pay gap outside academia. A similar pattern is observed in Sweden. We see that in Sweden, women in academia earn on average 11 percent less than men. Put differently, female PhDs in Sweden earn 89 kroner per 100 kroner a male PhD in academia earn. The pay gap in Norway is very similar inside academia.

When we adjust for observable characteristics, the gender pay gap is reduced. This means that some of the observed pay gap is attributable to observable differences in characteristics that influence wages – for instance gender differences in field of study, position, occupation or industry. We will explore the sources of the gap in more detail in the decomposition analysis. Still, when we control for observable differences among male and female PhDs, a significant

gender pay gap remains. Inside academia, the gender pay gap is very similar in Sweden and Norway, but there are clear country differences in the gender pay gap outside academia.

We also provide a robustness check on the role of full-time employment for our results. The gender pay gap is reduced further when we include this as a control, but the overall conclusion does not change.

Next, we proceed with decomposing the above presented gender pay gap in Table 4. Note that the coefficients are reported in terms of men's higher pay than women's – the opposite of the estimated coefficients in table 3.

From the decomposition it is clear that a large share of the gender pay gap is unexplained in the model. A general result is that up to 50 percent of the gap is explained in our models, but for the most part, less than 50 percent is explained by differences in observable characteristics. When we look at the detailed decomposition, it is clear that gender differences in individual characteristics, such as age, and origin explain very little of the gender pay gap. The same holds for gender differences in family characteristics, such as the marriage rate and number of children men and women have. Gender segregation in job-characteristics: position, occupation and industry, is accounts for a large share of the pay gap in most specification. Interestingly, the educational characteristics of female PhDs in Sweden pushes women's wage level up on average. Said differently; the gender pay gap would have been even higher had women had the same educational background as male PhDs.

The sum of the explained part from educational- and job-characteristics, show that gender segregation, meaning that men and women work in different fields, in different occupations and different occupations, account for a large share of the explained gender pay gap, and in some specifications, segregation accounts for more than the observed gap in levels.

Table 4: Oaxaca decomposition of the pay gap. Detailed decompositions. Sweden and Norway (2017).

	Sweden			Norway		
	All	Inside academia	Outside academia	All	Inside academia	Outside academia
Decomposition:						
Gender pay gap	0.167*** (0.0007)	0.119*** (0.006)	0.196*** (0.010)	0.223*** (0.006)	0.102*** (0.008)	0.285*** (0.005)
Explained	0.048*** (0.004)	0.059*** (0.004)	0.059*** (0.007)	0.054*** (0.004)	0.041*** (0.005)	0.092*** (0.006)
Unexplained	0.119*** (0.006)	0.059*** (0.005)	0.136*** (0.009)	0.169*** (0.005)	0.061 (0.007)	0.193*** (0.007)
Detailed decomposition:						
Explained						
Individual characteristics	- 0.004** (0.002)	-0.003*** (0.001)	0.0002 (0.003)	0.014*** (0.002)	-0.003 (0.0022)	0.018*** (0.002)
Family Characteristics	0.003*** (0.001)	0.002*** (0.001)	0.003** (0.001)	0.004*** (0.0007)	0.0018** (0.008)	0.004*** (0.0008)
Job characteristics	0.003*** (0.001)	0.078*** (0.004)	0.063*** (0.006)	0.025*** (0.002)	0.033*** (0.003)	0.062*** (0.004)
Education characteristics	-0.013*** (0.003)	-0.017*** (0.002)	-0.007 (0.004)	0.011** (0.002)	0.009** (0.004)	0.007** (0.003)
Unexplained						
Individual characteristics	-0.863*** (0.185)	-0.636*** (0.133)	-1.087*** (0.276)	-0.33** (0.140)	-0.186 (0.186)	0.088 (0.179)
Family Characteristics	0.053*** (0.009)	0.021*** (0.006)	0.068*** (0.015)	0.002 (0.019)	-0.035 (0.026)	-0.00004 (0.0237)
Job characteristics	0.136*** (0.021)	0.039*** (0.009)	0.262*** (0.030)	0.032 (0.035)	-0.005 (0.004)	0.062 (0.054)
Education characteristics	-0.009 (0.041)	0.019 (0.019)	-0.187* (0.110)	0.0039 (0.037)	0.061 (0.041)	-0.065 (0.040)
Constant	0.802*** (0.185)	0.616 *** (0.131)	1.08*** (0.288)	0.462*** (0.149)	0.227 (0.194)	0.108 (0.191)

Note: Standard errors in parenthesis. The three decompositions include the following sets of control variables; *All*: Individual= age, age², born in Norway/Swedish citizen. Family= married, children (dummy variable), Education= years since graduation (dummy variable), field of study (dummy variable), Job= academia (dummy variable), industry (dummy variable). *Inside academia*: Individual= age, age², born in Norway/Swedish citizen. Family= married, children (dummy variable), Education= years since graduation (dummy variable), field of study (dummy variable), Job=position (dummy variable). *Outside Academia*: Individual= age, age², born in Norway/Swedish citizen. Family= married, children (dummy variable), Education= years since graduation (dummy variable), field of study (dummy variable), Job= industry (dummy variable), occupation (dummy variable). * p<0.1, ** p<0.05, *** p<0.01.

Exploring the parenthood pay penalty

During the child-rearing years, the gender pay gap commonly is exacerbated since women tend to take on most of the care responsibilities. However, highly educated men commonly take on more childcare responsibilities compared to less educated men. We approach the question on whether having children leads to a pay penalty by estimating equation (2) separate for Sweden and Norway, and for all, inside academia and outside academia.

The point estimates from the regressions are reported in table 5. Output from regressions with interactions produce estimates that are interpreted relative to a reference group. Therefore, it can be tedious to compare the pay gap between mother's/father's and comparable women/men without children. We therefore compute linear combinations of coefficients⁷ for women and men separately and present them in figures 1-4.

Table 5. The role of children: Point estimates from separate regressions.

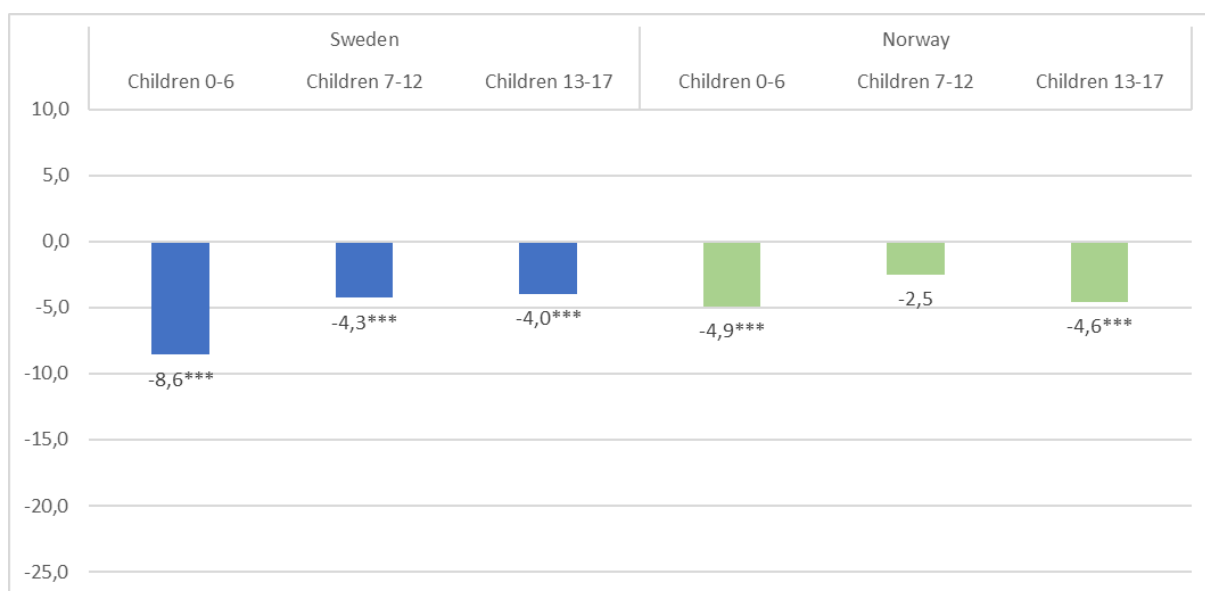
	Sweden			Norway		
	All	Inside academia	Outside academia	All	Inside academia	Outside academia
Child aged 0-5						
Woman	-0.103*** (0.007)	-0.039*** (0.005)	-0.124*** (0.010)	-0.161*** (0.006)	-0.053*** (0.008)	-0.175*** (0.007)
Child	-0.078*** (0.010)	-0.050*** (0.007)	-0.090*** (0.015)	0.007 (0.010)	0.003 (0.015)	-0.011 (0.013)
Woman# Child	-0.101*** (0.014)	-0.096*** (0.011)	-0.090*** (0.020)	-0.064*** (0.015)	-0.063*** (0.022)	-0.056*** (0.018)
Constant	4.85*** (0.092)	7.12*** (0.065)	4.06*** (0.142)	11.25*** (0.066)	12.08*** (0.089)	11.52*** (0.096)
Child aged 6-12						
Woman	-0.118*** (0.007)	-0.058*** (0.005)	-0.139*** (0.010)	-0.162*** (0.006)	-0.052*** (0.008)	-0.179*** (0.007)
Child	0.052*** (0.010)	0.014* (0.007)	0.054*** (0.015)	0.019* (0.011)	0.026* (0.015)	0.011 (0.014)
Woman# Child	-0.024*** (0.014)	0.004 (0.011)	-0.025 (0.021)	-0.061*** (0.016)	-0.079*** (0.022)	-0.032 (0.019)
Constant	4.64*** (0.092)	6.94*** (0.065)	3.84*** (0.143)	11.20*** (0.066)	12.03*** (0.089)	11.44*** (0.093)
Child aged 13-17						
Woman	-0.124*** (0.006)	-0.059*** (0.005)	-0.145*** (0.009)	-0.174*** (0.006)	-0.063*** (0.008)	-0.185*** (0.007)
Child	0.048*** (0.013)	0.019** (0.009)	0.052*** (0.019)	0.024** (0.011)	0.016 (0.015)	0.037** (0.014)
Woman# Child	0.011 (0.019)	0.021 (0.014)	0.009 (0.028)	0.023 (0.017)	0.024 (0.021)	0.022 (0.021)
Constant	4.64*** (0.092)	6.96*** (0.065)	3.84*** (0.142)	11.27*** (0.067)	12.09*** (0.092)	11.49*** (0.093)

Note: All regressions include the full set of control variables from table 2. Standard errors in parenthesis. *p<0.1, ** p<0.05, *** p<0.01.

⁷ Using the command `lincom` in STATA.

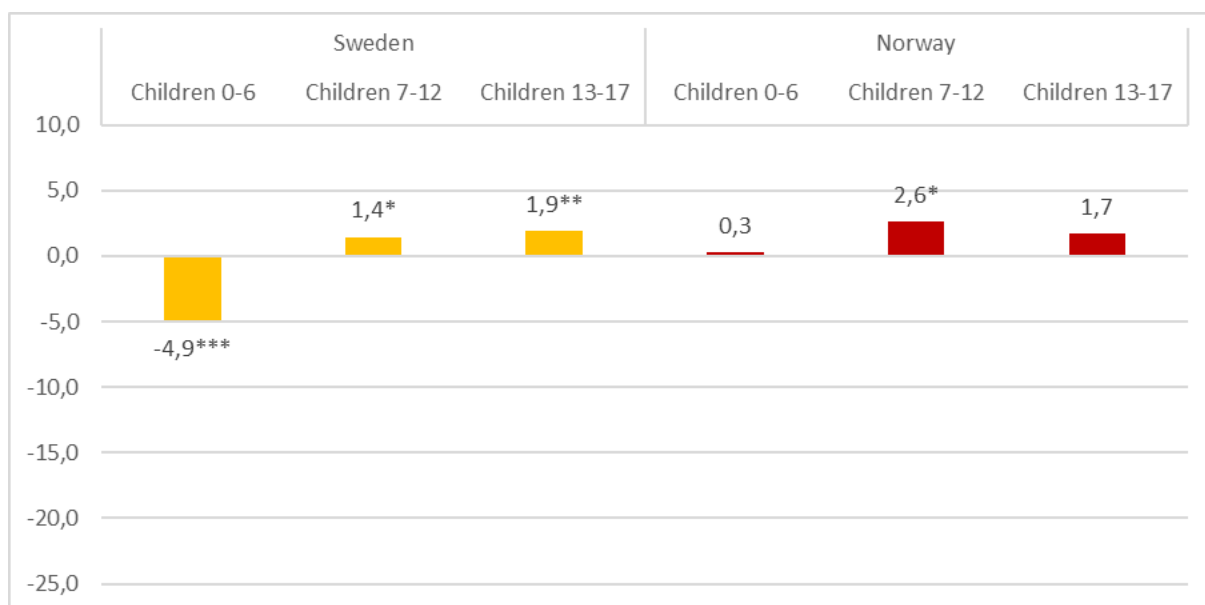
As seen in figure 1, there is a clear motherhood penalty in pay inside academia in both Sweden and Norway. The size of the gap is larger the smaller children the mother has. For men, the pattern is less clear. Interestingly, in figure 2 there is a small pay premium for men with older children inside academia, but in Sweden, men with young children also face a pay penalty relative to comparable men without children.

Figure 1. The motherhood penalty – Inside academia. Sweden and Norway (2017)



*p<0.1, ** p<0.05, *** p<0.01.

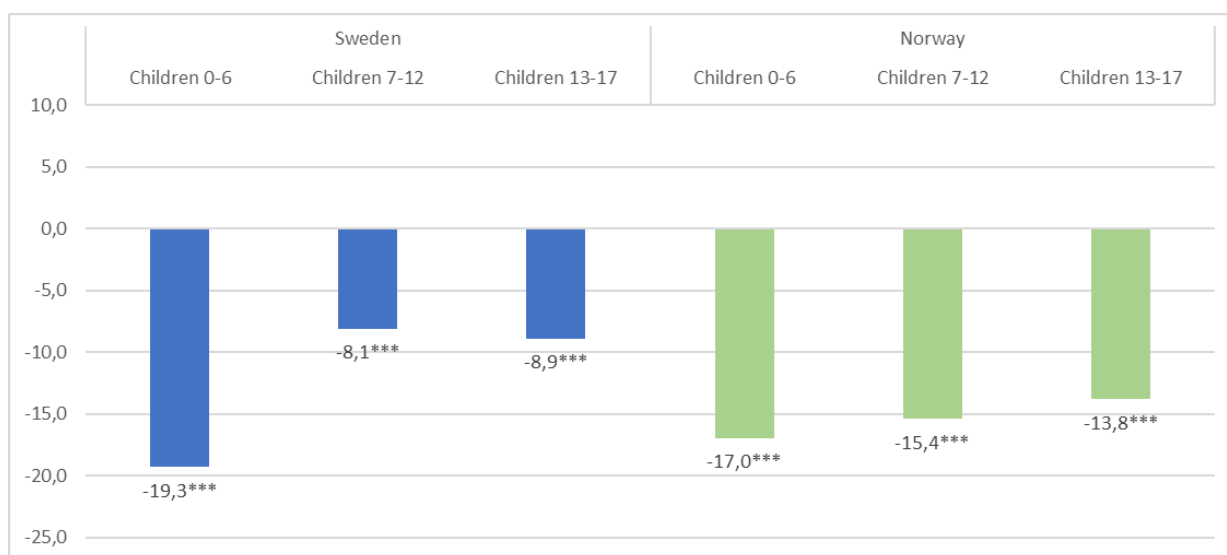
Figure 2. The fatherhood premium – Inside academia. Sweden and Norway (2017)



*p<0.1, ** p<0.05, *** p<0.01.

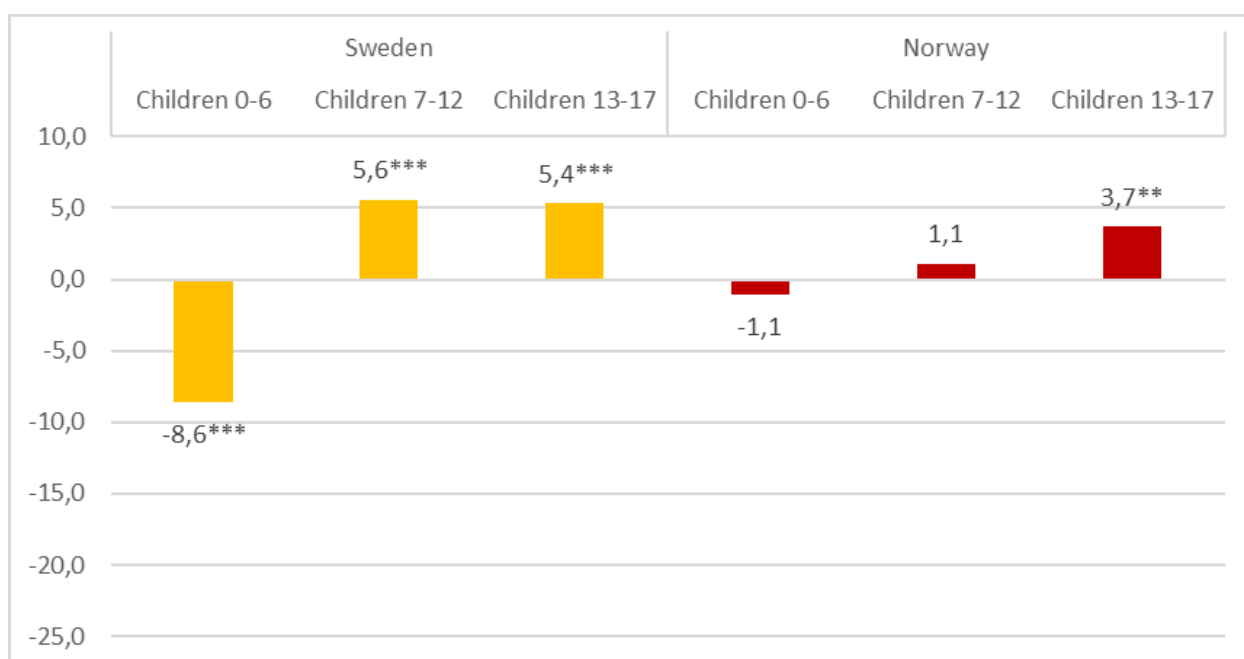
The motherhood penalty is largest outside academia. In Sweden, there is a slightly lower penalty for women with older children, but the gap remains relatively high across child age in Norway. Among men, the pattern remains, and men with young children face a penalty in Sweden. There are not clear signs of a similar pattern for Norway although: the point estimate for children aged 0-6 is negative, but not statistically significant.

Figure 3. The fatherhood premium – Outside academia. Sweden and Norway (2017)



*p<0.1, ** p<0.05, *** p<0.01.

Figure 4. The fatherhood premium – Outside academia. Sweden and Norway (2017)



*p<0.1, ** p<0.05, *** p<0.01.

References

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 436-455.

Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693-709.

Appendix

Table A1: Robustness check Including research institutes inside academia. Norway (2017).

Norway			
	All	Inside academia	Outside academia
Unadjusted	-0,223*** (0,006)	-0,120*** (0,006)	-0,329*** (0,010)
Adjusted	-0,151*** (0,005)	-0,073*** (0,006)	-0,208*** (0,008)
Observations	26 983	13678	13 295

Note: Standard errors in parenthesis. The three models include the following sets of control variables; *All*: female, age, age²born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), industry (dummy variable). *Inside academia*: female, age, age²born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), position (dummy variable). *Academia*: female, age, age²born in Norway/Swedish citizen, married, children (dummy variable), years since graduation (dummy variable), field of study (dummy variable), industry (dummy variable), occupation (dummy variable). * p<0.1, ** p<0.05, *** p<0.01.