An emergency room surgeon is set to operate on a boy who has been critically injured in a car accident in which his father was instantly killed. The surgeon takes one look at the boy's face and says, "I can't operate; he is my son."

The riddle is, how can this be?

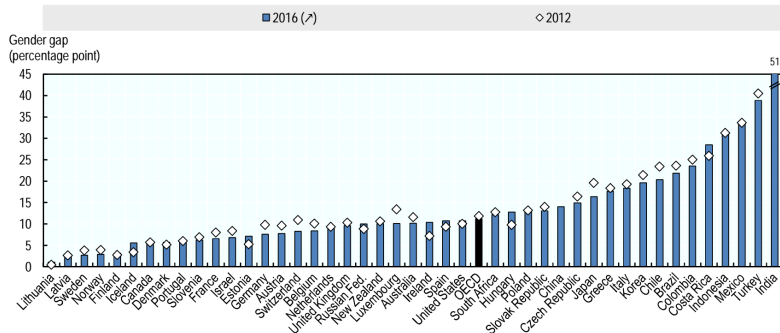# Decomposing the Gender Wage Gap

Daniela Piazzalunga

University of Verona, CHILD-CCA, IZA

XVIII Winter School on Inequality and Social Welfare Theory

10 Jan 2018

Gender differences in the labour market

Women experience lower employment rates

Gender gaps in employment rates, 2012 and 2016

Gender gap (M-F) in employment rate, 15-64 y.o.

Source: OECD, 2017

See Olivetti and Petrongolo 2016 for longer periods

Gender gap in employment

- The gap has narrowed in the postwar period
- Higher female participation lab mkt thanks to
  - More time freed up from household tasks
    - contraceptive pill
    - domestic technology
    - childcare availability
  - rising real wages for women (pull factor)
  - demand factor: rise of service sector
  - changing social norms (direction of causation?)
- Still, persistence of the gap

Unemployment rates

It depends on the country

- No difference in the US
- Converging also in Europe
- Still higher unemployment rates among women in southern Europe (eg Italy, Greece)
    - F low attachment to the labour market » more at risk of unemployment
    - institutional factors:
      more compressed wage distribution, reduction of labour turnover, temporary contracts increase the unemployment rate of workers less attached to the labour market

Women work less hours per week

- Part-time work
  - Large differences across countries, with the largest percentage in Netherlands (75%), Switzerland (66%), Germany (49%)
  - Increases women participation to the labour market
  - May have adverse effects on career and wages (per month, but also per hour)

- Involuntary part-time is a problem
  e.g. in Italy between 35% (2007; pre-crisis) and 60% (2016) of female part-timers are involuntary

Women work in different occupations and industries

- Unequal distribution of men and women across occupations, firms, industry
  - linked to persistent differences in the choice of field of study
- Female overrepresentation in the public sector
- Vertical segregation/glass ceiling
  - women promoted at lower rates than men

Gender wage gap

The unadjusted gender wage gap

Difference between the men's wage and women's wage divided by the men's wage.

$$\Delta \overline{Y} = \frac{\overline{Y}_M - \overline{Y}_F}{\overline{Y}_M}$$
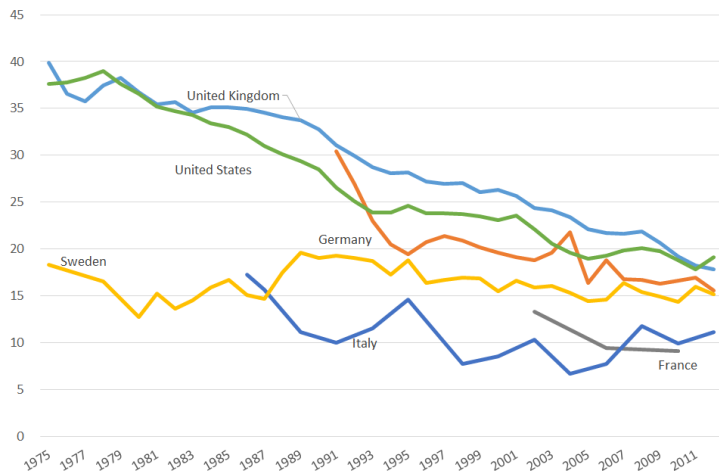
Useful as an indicator easy to read and intuitive;
often used by international institutions.

EU gap per hour 16% (Eurostat 2015, hourly data)

Issues with the definition

- does wage include overtime, bonuses, ...?
- monthly vs. hourly wages: pros and cons
- what about self-employed?

Gender wage gap
selected OECD countries, 1975-2012

Source: OECD, 2017

Gender wage gap:
trend over time

Narrowing...

- thanks to increase in
  - women education » reversal of the gap
  - women experience
- but at different speeds over the period and across countries
- more slowly since 1990

- in Italy (and few other countries) it increased during the economic crisis

Gender wage gap:
differences among countries

- degree of inequality in national wage structure
  - a more compressed wage structure reduce the gwg
- employment rate
- part-time work
- unemployment rate
- gender segregation
- family-friendly policies
  - e.g. subsidies, parental leave, childcare

Issues with comparability

OECD data:

- one of the largest scope & time coverage
- but based on national sources and definitions
- median wage (usually: mean wage)
- annual, monthly or weekly wages
- only full-time workers

More generally:

- self-reported vs. admin data
- some data (e.g. SES, used by Eurostat) exclude some sectors

Why do women earn less than men?

The gender pay gap is due to **structural factors** and to **discrimination**.

Structural factors:

- Different investment in human capital (much reduced):
  education, work experience, ...
  Persisting differences in college majors

- Occupational sectors
  (choice or noy?)

- Different preferences and behaviours (non cognitive skills):
  e.g. bargaining, risk aversion, ...
  Account for a smaller portion than occupation

Which in turn they can be a consequence of previous discrimination, cultural norms, etc...

Discrimination

"Individuals identical as regard their productive ability are treated differently because of certain of their nonproductive characteristics such as race, gender, or sexual orientation" (Cahuc et al., 2014)

- Pre-labour market discrimination;
- Labour market discrimination
    - Segregation:
      vertical or horizontal.

    - Wage discrimination
        - Taste discrimination;
          Employer, colleagues, consumers. (Becker, 1957)
        - Statistical discrimination;
          (Phelps, 1972)
        - Implicit discrimination
          (Bertrand et al., 2005)
          https://implicit.harvard.edu/implicit/

Analysing the gender wage gap:
Oaxaca-Blinder decomposition

Separate linear regressions

$$\ln \overline{Y}_g = \overline{X}_g \hat{\beta}_g$$

$\ln Y_g = \log$ (hourly) wage

$g = $ male, female

X: socio-economic characteristics, which (correctly?) influence wages

Separate linear regressions

$$\ln \overline{Y}_g = \overline{X}_g \hat{\beta}_g$$

$\ln Y_g = \log$ (hourly) wage

$g = $ male, female

X: socio-economic characteristics, which (correctly?) influence wages

Age, residence, education, work experience, sectors, professions, part-time, marital status...
Debate about what factors to include

Wage equation, separated for men and women

Men and women have different qualifications,
which may "justify" why male wages are larger.
E.g. experience.

On top of that, men and women may have different
returns to the same qualifications/characteristics.

Estimating separately men and women wage equation
allows us to estimate the different returns.

Oaxaca-Blinder decomposition

$$
\begin{aligned}
GWG = \ln \overline{Y}_m - \ln \overline{Y}_f &= \\
&= \overline{X}_m \hat{\beta}_m - \overline{X}_f \hat{\beta}_f = \\
&= \overline{X}_m \hat{\beta}_m - \overline{X}_f \hat{\beta}_f + (\overline{X}_f \hat{\beta}_m - \overline{X}_f \hat{\beta}_m) = \\
&= \underbrace{(\overline{X}_m - \overline{X}_f)\hat{\beta}_m}_{\substack{\text{Wage gap due to} \\ \text{different characteristics} \\ \text{('explained')}}} + \underbrace{\overline{X}_f(\hat{\beta}_m - \hat{\beta}_f)}_{\substack{\text{Wage gap due to} \\ \text{different returns} \\ \text{('unexplained')}}}
\end{aligned}
$$

It also allow to assess the contribution of different factors (detailed decomposition)

Limitations

- The unexplained component could be overestimated or underestimated (omitted variables, self-selection)
- What influence control variable?
  Discrimination, culture, preferences, institutions
- No understanding of the underling mechanism
- Generally, we cannot give a causal interpretation of the resuls
- Partial equilibrium approach

Issues

1. Counterfactual choice
2. Self-selection
   - Heckman correction
   - Imputation

Issue 1: Alternative choices of counterfactual I

The decomposition above makes use of $\overline{X}_f \hat{\beta}_m$, the counterfactual wage of women should they have the wage structure of men.

However, the male wage structure may not represent the appropriate counterfactual for the way women would be paid in the absence of labor market discrimination.

Issue 1: Alternative choices of counterfactual II

Several alternatives:

- (Standard) Male wage structure:
  $$GWG = \underbrace{(\overline{X}_m - \overline{X}_f)\hat{\beta}_m}_{\text{'explained'}} + \underbrace{\overline{X}_f(\hat{\beta}_m - \hat{\beta}_f)}_{\text{'unexplained'}}$$

- Female wage structure (typically, explained part smaller)

  $$GWG = \underbrace{(\overline{X}_m - \overline{X}_f)\hat{\beta}_f}_{\text{'explained'}} + \underbrace{\overline{X}_m(\hat{\beta}_m - \hat{\beta}_f)}_{\text{'unexplained'}}$$

- Weighted average of the two

Issue 1: Alternative choices of counterfactual III

- Pooled regression w/o group-specific intercept (Neumark 1988)

  Problematic: coefficients biased due to omission of the group intercept » Overestimated role of observables.

- Pooled regression with group intercept
  Unexplained component usually in between male and female 'approach'
  Detailed decomposition based on triple difference

$$GWG = \underbrace{(\overline{X}_m - \overline{X}_f)\hat{\beta}^*}_{(\text{'explained'})} + \underbrace{\overline{X}_m(\hat{\beta}_m - \hat{\beta}^*) + \overline{X}_f(\hat{\beta}^* - \hat{\beta}_f)}_{(\text{'unexplained'})}$$

Source: Elder et al. 2010; Jann 2008; Fortin et al. 2011

Issue 2: Selection into employment

- Gender wage gap (usually) negatively correlated with gender gap in employment
- The fewer women employed, the more they are positively selected and the higher their realtive wage
- Not a uniform relation: this is not the case if we compare Italy vs. Japan, but also looking at Italian regions (Southern Italy: larger gap in employment but also larger gwg).

Issue 2: Selection into employment

- Heckman correction (1974, 1976):
  introduce into the wage equation a correction term obtained
  from a first model on the probability of being employed
  - Issue: exclusion restrictions (e.g. number of children,
    nonlabour income) could be argued to directly affect wages

  Blau and Kahn 2017

- Median gaps & imputing wages
  e.g. for Italy median gwg 1999 from 5.9% up to 23%
  - Issue: assumptions needed

  Olivetti and Petrongolo 2008

Additional issues

3. Omitted group problem: an issue in detailed decompositions with categorical variables for the unexplained part
   - Express effects as deviation from the grand mean

4. Support: owing to gender-related pre-market sorting, there are many women for whom there are no convincing male comparables.
   Non-parametric approach - Black et al 2008
   - Matching rates for college graduates about 87%
   - Issues with experience comparability (include only half of female sample)
   - parametric estimates are larger (small differences for white women, striking differences for minorities)

An application of the OB decomposition

Schooling and work experience by gender, US 1981-2011

Year	Men	Women	Difference: Men–Women
Years of schooling			
1981	13.3	13.2	0.2
1990	13.8	13.7	0.0
1999	14.2	14.3	–0.1
2011	14.3	14.5	–0.2
Bachelor's degree only			
1981	18.1%	15.3%	2.7%
1990	20.0%	17.6%	2.3%
1999	23.4%	22.2%	1.2%
2011	26.2%	24.7%	1.5%
Advanced degree			
1981	10.0%	7.4%	2.5%
1990	10.3%	8.7%	1.6%
1999	11.7%	10.8%	0.9%
2011	12.9%	15.7%	–2.8%
Years of full time experience			
1981	20.3	13.5	6.8
1990	19.2	14.7	4.5
1999	19.8	15.9	3.8
2011	17.8	16.4	1.4

Source: Blau and Kahn, 2017

Incidence of managerial or professional jobs, US 1981-2011

Year	Men	Women	Difference: Men–Women
Managerial jobs			
1981	21.5%	9.2%	12.3%
1990	21.1%	10.9%	10.2%
1999	21.8%	15.3%	6.5%
2011	18.3%	16.2%	2.2%
Professional jobs			
1981	17.0%	21.8%	–4.8%
1990	19.4%	26.1%	–6.6%
1999	20.4%	26.9%	–6.4%
2011	21.7%	31.1%	–9.4%
"Male" professional jobs			
1981	14.6%	10.1%	4.5%
1990	17.3%	14.1%	3.2%
1999	17.6%	13.2%	4.4%
2011	18.6%	17.8%	0.8%

Source: Blau and Kahn, 2017

Decomposition of Gender wage gap, US 1980 and 2010

| | 1980 | | 2010 | |
| | Effect of gender gap in explanatory variables | | Effect of gender gap in explanatory variables | |
Variables	log points	Percent of gender gap explained	log points	Percent of gender gap explained
Panel A. Human-capital specification				
Education variables	0.0129	2.7	–0.0185	–7.9
Experience variables	0.1141	23.9	0.0370	15.9
Region variables	0.0019	0.4	0.0003	0.1
Race variables	0.0076	1.6	0.0153	6.6
Total explained	0.1365	28.6	0.0342	14.8
Total unexplained gap	0.3405	71.4	0.1972	85.2
Total pay gap	0.4770	100.0	0.2314	100.0
Panel B. Full specification				
Education variables	0.0123	2.6	–0.0137	–5.9
Experience variables	0.1005	21.1	0.0325	14.1
Region variables	0.0001	0.0	0.0008	0.3
Race variables	0.0067	1.4	0.0099	4.3
Unionization	0.0298	6.2	–0.0030	–1.3
Industry variables	0.0457	9.6	0.0407	17.6
Occupation variables	0.0509	10.7	0.0762	32.9
Total explained	0.2459	51.5	0.1434	62.0
Total unexplained gap	0.2312	48.5	0.0880	38.0
Total pay gap	0.4770	100.0	0.2314	100.0

Source: Blau and Kahn, 2017

Substantial role for occupation and industry

The role of psychological traits

Study	Sample	Traits examined	Raw gender wage gap (logs)	Effect of gender differences in psych. factors on gender pay gap (logs)	Percentage of gender pay gap due to gender differences in psych. traits
Mueller and Plug (2006)	Wisconsin 1957 HS grads, 1992 data	"Big 5": extroversion; agreeableness; conscientiousness; neuroticism; openness	0.587	0.043–0.095	7.3–16.2
Semykina and Linz (2007)	Russia 2000–2003	Locus of control; challenge/affiliation	0.311–0.397	0.012–0.026	3.0–8.4
Fortin (2008)	US NELS 1972 and 1988 cohorts: 1979, 1986, and 2000	Self-esteem; locus of control; money/ work importance; people/family importance	0.181–0.237	0.008–0.032	4.4–14.0
Manning and Swaffield (2008)	British cohort study: 1970 birth cohort, 2000 data	Risk; competitiveness; self-esteem; other-regarding; career orientation; locus of control	0.203	0.005–0.056	2.5–27.6
Nyhus and Pons (2012)	Netherlands 2005	Locus of control; time preference	0.246	0.028–0.035	11.5–14.1
Reuben, Sapienza, and Zingales (2015)	2008 Univ. of Chicago Booth MBA cohort	Taste for competition	0.119	0.010–0.012	8.4–10.1
Cattan (2014)	NLSY 1979, 4 points in life cycle	Self-confidence	0.18–0.30	0.010–0.036	5.4–14.5

Possible extensions of the OB decomposition

- Gender wage gap over time
- Double decomposition;
- Beyond the mean: quantile decomposition

Change in wages over time for a given group

The standard OB decomposition can be used to decompose ΔY over time for a given group g:

$$\Delta Y_g = ln\overline{Y}_g^1 - ln\overline{Y}_g^0$$
$$= (\overline{X}_g^1 - \overline{X}_g^0)\hat{\beta}_g^0 + \overline{X}_g^1(\hat{\beta}_g^1 - \hat{\beta}_g^1)$$

In this example, using $t = 0$ as counterfactual.

Gender wage gap over time
Extended Oaxaca-Blinder decomposition

How much of the change in the gender wage gap is due to changes in individual characteristics and how much to changes in the wage structures.

$$\Delta GWG = \Delta \ln \overline{Y}_m - \Delta \ln \overline{Y}_f$$
$$= (\ln \overline{Y}_m^1 - \ln \overline{Y}_m^0) - (\ln \overline{Y}_f^1 - \ln \overline{Y}_f^0)$$

$$= [\underbrace{(\overline{X}_m^1 - \overline{X}_m^0)\widehat{\beta}_m^0}_{\substack{\Delta \text{ M wages due to} \\ \text{changes in M char.}}} + \underbrace{(\widehat{\beta}_m^1 - \widehat{\beta}_m^{09})\overline{X}_m^1}_{\substack{\Delta \text{ M wages due to} \\ \text{changes in M return}}}] - [\underbrace{(\overline{X}_f^1 - \overline{X}_f^{09})\widehat{\beta}_f^0}_{\substack{\Delta \text{ F wages due to} \\ \text{changes in F char.}}} + (\underbrace{\widehat{\beta}_f^1 - \widehat{\beta}_f^0)\overline{X}_f^1}_{\substack{\Delta \text{ F wages due to} \\ \text{changes in F return}}}]$$

$$= [\underbrace{(\overline{X}_m^1 - \overline{X}_m^{09})\widehat{\beta}_m^0 - (\overline{X}_f^1 - \overline{X}_f^0)\widehat{\beta}_f^0}_{\substack{\Delta \text{ GWG due to} \\ \text{changes in char.}}}] + [\underbrace{(\widehat{\beta}_m^1 - \widehat{\beta}_m^{09})\overline{X}_m^1 - (\widehat{\beta}_f^1 - \widehat{\beta}_f^0)\overline{X}_f^1}_{\substack{\Delta \text{ GWG due to} \\ \text{changes in return}}}]$$

Kassenboehmer and Sinning 2014

Extended OB I: Changes in wages, by gender

Men	Actual wages and decomposition	
2011 (a)	2.35***	(0.01)
2009 (b)	2.39***	(0.01)
Change (c)	-0.04***	(0.01)
Due to Δ characteristics (d)	0.01**	(0.01)
Due to Δ return (e)	-0.05***	(0.01)
Women		
2011 (g)	2.28***	(0.01)
2009 (h)	2.33***	(0.01)
Change (i)	-0.05***	(0.01)
Due to Δ charact. (l)	0.01**	(0.01)
Due to Δ return (m)	-0.06***	(0.01)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.
Robust standard errors in parentheses. Log wages in 2008 real
prices.
Benchmark coefficients: 2009.

Extended OB II: Change in the gender wage gap

Gender Wage Gap	Actual wages and decomposition	
2011 (a)-(g)	0.07***	(0.01)
2009 (b)-(h)	0.06***	(0.01)
Δ GWG (c)-(i)	0.01	(0.01)
Total Δ characteristics (d)-(l)	-0.00	(0.01)
Total Δ return (e)-(m)	0.01*	(0.01)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.
Benchmark coefficients: 2009.

Double decomposition
An augmented OB decomposition

To estimate the double-negative effect of being (e.g.) a migrant and a woman at the same time (or black & woman in the U.S.)

The difference in mean log wages of a native-born (N) male (m) and a foreign-born (I) female (f) is:

$$\overline{lnY}_m^N - \overline{lnY}_f^I = \underbrace{(\overline{lnY}_m^N - \overline{lnY}_m^I)}_{\substack{\text{Native-migrant w gap} \\ \text{among men}}} + \underbrace{(\overline{lnY}_m^I - \overline{lnY}_f^I)}_{\substack{\text{Gender w gap} \\ \text{among migrants}}}$$

Also in this case, an alternative decomposition is possible:

$$= \underbrace{(\overline{lnY}_m^N - \overline{lnY}_f^N)}_{\substack{\text{Gender w gap} \\ \text{among natives}}} + \underbrace{(\overline{lnY}_f^N - \overline{lnY}_f^I)}_{\substack{\text{Native-migrant w gap} \\ \text{among women}}}$$

Source: Shamsuddin 1998

Double decomposition

Following the OB decomposition, the (first) total gap can be rewritten as:

$$\overline{lnY}_m^N - \overline{lnY}_f^I = (\overline{lnY}_m^N - \overline{lnY}_m^I) + (\overline{lnY}_m^I - \overline{lnY}_f^I)$$

$$= [(\overline{X}_m^N - \overline{X}_m^I)\hat{\beta}_m^N + (\hat{\beta}_m^N - \hat{\beta}_m^I)\overline{X}_m^I] + [(\overline{X}_m^I - \overline{X}_f^I)\hat{\beta}_m^I + (\hat{\beta}_m^I - \hat{\beta}_f^I)\overline{X}_f^I]$$

Double decomposition

(a) $(\overline{X}_m^N - \overline{X}_m^I)\hat{\beta}_m^N$: explained native-migrant gap within men

(b) $(\hat{\beta}_m^N - \hat{\beta}_m^I)\overline{X}_m^I$: unexplained native-migrant gap within men

$\qquad\qquad\qquad\qquad\qquad$ (native-migrant 'discrimination')

(c) $(\overline{X}_m^I - \overline{X}_f^I)\hat{\beta}_m^I$: explained gender wage gap within immigrants

(d) $(\hat{\beta}_m^I - \hat{\beta}_f^I)\overline{X}_f^I)$: unexplained gender wage gap within immigrants

$\qquad\qquad\qquad\qquad\qquad$ (gender 'discrimination')

The 'double discrimination' is the sum of (b) + (d)

The choice of the order of the decomposition affects the results ('path dependency').

Double-negative effect, Italy 2009

	Log wage gap	Percentage
Migrant gender gap	0.17	32%
Italian gender gap	0.06	10%
Men ethnic gap	0.36	68%
Women ethnic gap	0.48	90%
Double-negative effect	0.53	

Double-negative effect, Italy 2009

Double negative effect (difference in mean log wages) (1) + (2)					0.53	
		Gender gap				
		I specific.		**II specific.**		
Migrants	Explained (a)	0.06***	(0.01)	−0.07***	(0.01)	Italians
		11.80%		−13.17%		
	Unexplained (b)	0.11***	(0.02)	0.13***	(0.01)	
		20.81%		24.07%		
		Ethnic gap				
		I specific.		**II specific.**		
Males	Explained (c)	0.19***	(0.01)	0.25***	(0.01)	Female
		35.03%		47.15%		
	Unexplained (d)	0.17***	(0.01)	0.22***	(0.01)	
		32.36%		41.94%		
		Double 'discrimination'				
	(b) + (d)	0.28		0.35		
		53.16%		66.01%		

Source: Piazzalunga, 2015

Final comments

The GWG has additional negative consequences:

- Pension gap;
- More women at risk of poverty;
  Es. after divorce; when elderly.
- Less wealth;
- Smaller bargaining power in the household;
- Reduced autonomy from the husband;
  Could it affect also the possibility of divorcing?

Thank you