Estimating gender wage gap in the presence of efficiency wages

Joanna Tyrowicz
(joint work with Katarzyna Bech)

FAME | GRAPE, IAAEU, IZA and University of Warsaw

IAFFE @ ASSA, 2018

Estimating gender wage gap in the presence of efficiency wages

Joanna Tyrowicz
(joint work with Katarzyna Bech)

FAME | GRAPE, IAAEU, IZA and University of Warsaw

IAFFE @ ASSA, 2018

Men and women do not appear to be earning the same

- **Gender wage gaps**: roughly 10-25% penalty
  Weichselbaumer and Winter-Ebmer (2005), Blau and Kahn (2016)

Men and women do not appear to be earning the same

- **Gender wage gaps**: roughly 10-25% penalty
  Weichselbaumer and Winter-Ebmer (2005), Blau and Kahn (2016)
- Multiplicity of methods to address data and labor market imperfections
  reviewed by Fortin et al. (2011) and Tyrowicz et al (2017)

Men and women do not appear to be earning the same

- **Gender wage gaps**: roughly 10-25% penalty
  Weichselbaumer and Winter-Ebmer (2005), Blau and Kahn (2016)
- Multiplicity of methods to address data and labor market imperfections
  reviewed by Fortin et al. (2011) and Tyrowicz et al (2017)

BUT labor markets are **segmented**

- Standard estimates: 15-30% premium
  Krueger and Summers (1988), Konings and Walsh (1994)

Men and women do not appear to be earning the same

- **Gender wage gaps**: roughly 10-25% penalty
  Weichselbaumer and Winter-Ebmer (2005), Blau and Kahn (2016)
- Multiplicity of methods to address data and labor market imperfections
  reviewed by Fortin et al. (2011) and Tyrowicz et al (2017)

BUT labor markets are **segmented**

- Standard estimates: 15-30% premium
  Krueger and Summers (1988), Konings and Walsh (1994)
- individual productivity unobservable
  indirect identification, e.g. Weiss (1980) or case study e.g. Campbell (1993)

Men and women do not appear to be earning the same

- **Gender wage gaps**: roughly 10-25% penalty
  Weichselbaumer and Winter-Ebmer (2005), Blau and Kahn (2016)
- Multiplicity of methods to address data and labor market imperfections
  reviewed by Fortin et al. (2011) and Tyrowicz et al (2017)

BUT labor markets are **segmented**

- Standard estimates: 15-30% premium
  Krueger and Summers (1988), Konings and Walsh (1994)
- individual productivity unobservable
  indirect identification, e.g. Weiss (1980) or case study e.g. Campbell (1993)

Men and women do not appear to be earning the same

Question

Can labor market segmentation be **gendered**?

Men and women do not appear to be earning the same

Question

Can labor market segmentation be **gendered**?

If so, **how much?**

Men and women do not appear to be earning the same

Question

Can labor market segmentation be **gendered**?

If so, **how much?**

And how does it **bias estimates** of gender wage gaps

Men and women do not appear to be earning the same

Question

Can labor market segmentation be **gendered**?

If so, **how much?**

And how does it **bias estimates** of gender wage gaps

We are not the first: Bulow (1986!) proposes **efficiency wages** as an explanation for GWG

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.

Motivation

How efficiency wages and GWG interact?

$\rightarrow$ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.
- If efficiency wages are selective then even adjusted GWG will confound

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.

- If efficiency wages are selective then even adjusted GWG will confound
  - below productivity compensating of women
    with

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.
- If efficiency wages are selective then even adjusted GWG will confound
  - below productivity compensating of women
    with
  - above productivity efficiency wage prevalence.

GRAPE

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.
- If efficiency wages are selective then even adjusted GWG will confound
  - below productivity compensating of women
    with
  - above productivity efficiency wage prevalence.

GRAPE

Motivation

How efficiency wages and GWG interact?

→ If men receive efficiency wages more often, (pooled) GWG estimates are **biased**

- Gender wage gap: women are paid *unjustifiably* less than men.
- Efficiency wages: a group of workers is paid *in excess of productivity*.
- If efficiency wages are selective then even adjusted GWG will confound
  - below productivity compensating of women
    with
  - above productivity efficiency wage prevalence.

  Selectivity: efficiency wages used more often in occupations and/or industries dominated by men.

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

- which separates workers into privileged and standard markets

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

- which separates workers into privileged and standard markets
- when separation is endogenous

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

- which separates workers into privileged and standard markets
- when separation is endogenous

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

- which separates workers into privileged and standard markets
- when separation is endogenous and unobservable

What we do

Contribution

We propose a new estimator of the adjusted gender wage gaps

- which separates workers into privileged and standard markets
- when separation is endogenous and unobservable

We apply our estimator to the EU countries (linked employer-employee data)

What we do

Contribution

Preview of the results

- women experience barriers accessing the privileged market

What we do

Contribution

Preview of the results

- women experience barriers accessing the privileged market
- adjusted GWGs differ between the privileged and standard markets

What we do

Contribution

Preview of the results

- women experience barriers accessing the privileged market
- adjusted GWGs differ between the privileged and standard markets
- accounting for the efficiency wages, adjusted GWGs different than in the pooled estimation

How to model unknown and endogenous split

The model

$$Y_i = \left\{ \begin{array}{l} Y_{1,i} \text{ iff } Y^*_{s,i} > 0 \\ Y_{0,i} \text{ iff } Y^*_{s,i} \leq 0 \end{array} \right\} \text{with}$$

How to model unknown and endogenous split

The model

$$Y_i = \left\{ \begin{array}{l} Y_{1,i} \text{ iff } Y^*_{s,i} > 0 \\ Y_{0,i} \text{ iff } Y^*_{s,i} \leq 0 \end{array} \right\} \text{with}$$

$$\begin{aligned} Y_{1,i} &= X_i\beta_1 + u_{1,i} & \leftarrow \text{"privileged market"} \\ Y_{0,i} &= X_i\beta_0 + u_{0,i} & \leftarrow \text{"standard market"} \\ Y^*_{s,i} &= W_i\alpha - v_i & \leftarrow \text{ the "split" mechanism} \end{aligned}$$

How to model unknown and endogenous split

The model

$$Y_i = \left\{ \begin{array}{c} Y_{1,i} \text{ iff } Y^*_{s,i} > 0 \\ Y_{0,i} \text{ iff } Y^*_{s,i} \leq 0 \end{array} \right\} \text{with}$$

$$\begin{aligned} Y_{1,i} &= X_i\beta_1 + u_{1,i} & \leftarrow \text{“privileged market”} \\ Y_{0,i} &= X_i\beta_0 + u_{0,i} & \leftarrow \text{“standard market”} \\ Y^*_{s,i} &= W_i\alpha - v_i & \leftarrow \text{ the “split” mechanism} \end{aligned}$$

Disturbances are jointly normally distributed with mean 0 and covariance matrix

$$\begin{pmatrix} \sigma_1^2 & 0 & \sigma_{1v} \\ 0 & \sigma_0^2 & \sigma_{0v} \\ \sigma_{1v} & \sigma_{0v} & \sigma_v^2 \end{pmatrix}.$$

How to model unknown and endogenous split

The difficulty

- OLS + probit if disturbances were pairwise uncorrelated and if the sample separation was known, i.e.

$$I_i = \left\{ \begin{array}{c} 1 \text{ iff } Y_i = Y_{1,i} \\ 0 \text{ iff } Y_i = Y_{0,i} \end{array} \right\}$$

How to model unknown and endogenous split

The difficulty

- OLS + probit if disturbances were pairwise uncorrelated and if the sample separation was known, i.e.

$$I_i = \left\{ \begin{array}{c} 1 \text{ iff } Y_i = Y_{1,i} \\ 0 \text{ iff } Y_i = Y_{0,i} \end{array} \right\}$$

- Endogenous switching regression if disturbances are correlated, but the sample split is known (e.g. `-movestay-`)

How to model unknown and endogenous split

The difficulty

- OLS + probit if disturbances were pairwise uncorrelated and if the sample separation was known, i.e.

$$I_i = \left\{ \begin{array}{c} 1 \text{ iff } Y_i = Y_{1,i} \\ 0 \text{ iff } Y_i = Y_{0,i} \end{array} \right\}$$

- Endogenous switching regression if disturbances are correlated, but the sample split is known (e.g. `-movestay-`)

How to model unknown and endogenous split

The difficulty

- OLS + probit if disturbances were pairwise uncorrelated and if the sample separation was known, i.e.

$$I_i = \left\{ \begin{array}{c} 1 \text{ iff } Y_i = Y_{1,i} \\ 0 \text{ iff } Y_i = Y_{0,i} \end{array} \right\}$$

- Endogenous switching regression if disturbances are correlated, but the sample split is known (e.g. `-movestay-`)

Obtaining the sample split

or squeezing blood out of the stone

Endogenous Switching Regression with an unknown sample separation

- Neumark and Wascher (1994, ILR) and Hovakimian and Titman (2006, JMC&B)

Obtaining the sample split

or squeezing blood out of the stone

Endogenous Switching Regression with an unknown sample separation

- Neumark and Wascher (1994, ILR) and Hovakimian and Titman (2006, JMC&B)
- expectation maximization algorithm (Dempster et al. 1977; Hartley 1978)

Obtaining the sample split

or squeezing blood out of the stone

Endogenous Switching Regression with an unknown sample separation

- Neumark and Wascher (1994, ILR) and Hovakimian and Titman (2006, JMC&B)
- expectation maximization algorithm (Dempster et al. 1977; Hartley 1978)

Obtaining the sample split

or squeezing blood out of the stone

Endogenous Switching Regression with an unknown sample separation

- Neumark and Wascher (1994, ILR) and Hovakimian and Titman (2006, JMC&B)
- expectation maximization algorithm (Dempster et al. 1977; Hartley 1978)

$$\ln L = \sum_{i=1}^{n} \left\{ (1 - I_i) \left[\ln \phi\left(\frac{u_{0,i}}{\sigma_0}\right) - \ln \sigma_0 + \ln \left\{ 1 - \Phi\left(\frac{W_i\alpha - \rho_0 \frac{u_{0,i}}{\sigma_0}}{\sqrt{1-\rho_0^2}} \right) \right\} \right] \right.$$
$$\left. + I_i \left[\ln \phi\left(\frac{u_{1,i}}{\sigma_1}\right) - \ln \sigma_1 + \ln \Phi\left(\frac{W_i\alpha - \rho_1 \frac{u_{1,i}}{\sigma_1}}{\sqrt{1-\rho_1^2}} \right) \right] \right\}$$

Obtaining the sample split

or squeezing blood out of the stone

Table: Variables determining split and determining wages

Variable	Switching regression W	Wage regression X
Age	Y	Y
Gender	Y	Y
Education	Y	Y
Occupation	Y	Y
Industry	Y	N

Obtaining the sample split

or squeezing blood out of the stone

Table: Variables determining split and determining wages

Variable	Switching regression W	Wage regression X
Age	Y	Y
Gender	Y	Y
Education	Y	Y
Occupation	Y	Y
Industry	Y	N

$+$ interactions between *gender* and all other variables.

Gender wage gap decomposition

After obtaining the estimates of the sample split

- We decompose GWG into six components:
  - explained and unexplained components from the switching equation
  - explained and unexplained components from the privileged market equation
  - explained and unexplained components from the standard market equation

Gender wage gap decomposition

After obtaining the estimates of the sample split

- We decompose GWG into six components:
  - explained and unexplained components from the switching equation
  - explained and unexplained components from the privileged market equation
  - explained and unexplained components from the standard market equation
- using Oaxaca-Blinder decomposition (any decomposition could be used!)

$$\ln \overline{W}_m - \ln \overline{W}_f = \beta^*(\overline{X}_m - \overline{X}_f) + \overline{X}_m(\beta_m - \beta^*) + \overline{X}_f(\beta^* - \beta_f).$$

- The choice of β^* following Słoczyński (2015).

Data

Structure of Earnings, Eurostat

- Linked employer-employee data
- *The largest* individual level data available (100k - 2m observations)
- Waves every two years
- Comparable methodology
- Sample design
  - All workers in small firms
  - Random selection of workers in medium and large firms
  - Only definition of small/medium/large varies across countries
- We use 2006 wave, all available countries (few dropped because of missing data)

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function ($I()$)

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function ($I()$) , which has no theoretical threshold

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function $(I())$, which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function ($I()$) , which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages
- One can pick other thresholds (below or above)

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function $(I())$, which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages
- One can pick other thresholds (below or above)
- Follow data: Cramer approach (predicted allocations to privileged market)

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function $(I())$, which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages
- One can pick other thresholds (below or above)
- Follow data: Cramer approach (predicted allocations to privileged market)

Table: Sample results - Poland

	OLS		Privileged market		Standard market		Switching	
Split	Raw	Adj.	Raw	Adj.	Raw	Adjusted	Raw	Adj.
85th	5.0%	23.6%	-51.8%	28.3%	13.8%	8.3%	3.9%	6.9%

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function $(I())$, which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages
- One can pick other thresholds (below or above)
- Follow data: Cramer approach (predicted allocations to privileged market)

Table: Sample results - Poland

	OLS		Privileged market		Standard market		Switching	
Split	Raw	Adj.	Raw	Adj.	Raw	Adjusted	Raw	Adj.
85th	5.0%	23.6%	-51.8%	28.3%	13.8%	8.3%	3.9%	6.9%
75th	5.0%	23.6%	-46.7%	27.4%	21.6%	8.1%	3.9%	6.9%
95th	5.0%	23.6%	.	.	1.8%	7.7%	3.9%	6.9%

Splitting the data before obtaining GWG estimates

Where is the "delineation" between privileged and standard markets?

- The estimated indicator function $(I())$, which has no theoretical threshold
- In empirical literature, typically 15% of workers receive efficiency wages
- One can pick other thresholds (below or above)
- Follow data: Cramer approach (predicted allocations to privileged market)

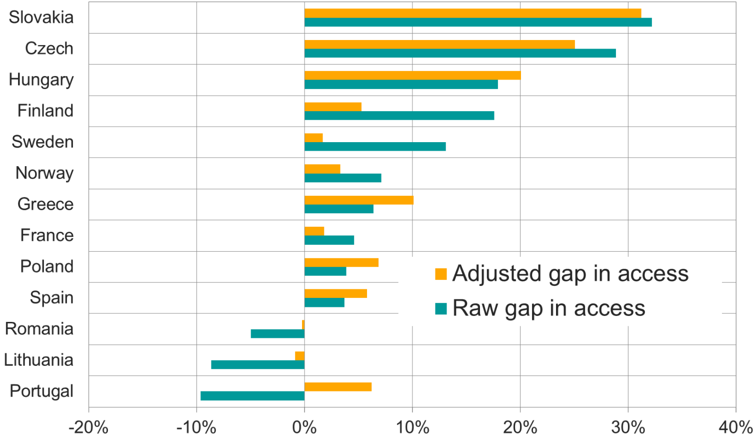
Table: Sample results - Poland

	OLS		Privileged market		Standard market		Switching	
Split	Raw	Adj.	Raw	Adj.	Raw	Adjusted	Raw	Adj.
85th	5.0%	23.6%	-51.8%	28.3%	13.8%	8.3%	3.9%	6.9%
75th	5.0%	23.6%	-46.7%	27.4%	21.6%	8.1%	3.9%	6.9%
95th	5.0%	23.6%	.	.	1.8%	7.7%	3.9%	6.9%
Cramer	5.0%	23.6%	-23.7%	26.8%	11.6%	7.8%	3.9%	6.9%

* Cramer at 42%

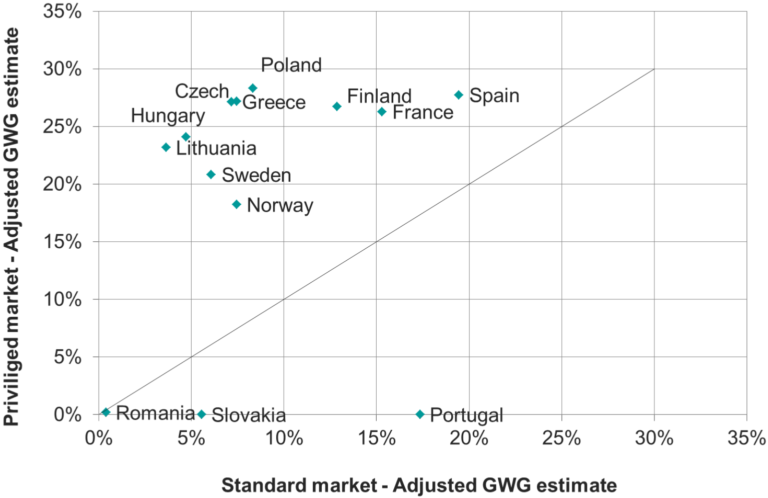
Women experience barriers accessing the privileged market

Switching regression decomposition – raw and adjusted gaps (LPM), 85% split



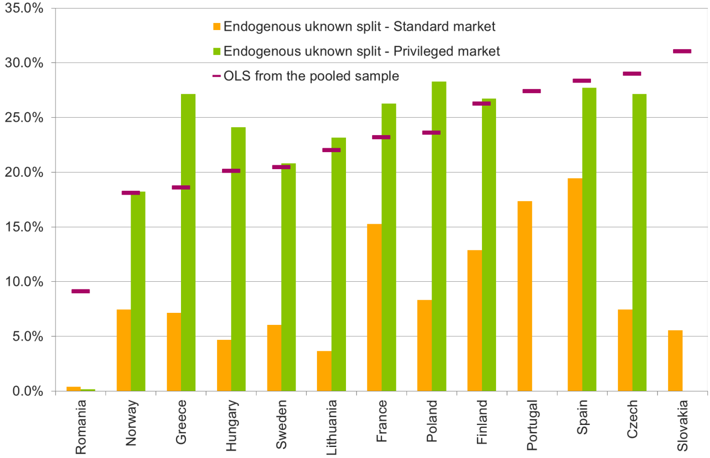
Adjusted GWGs differ between the markets

Scatter plot of the standard vs privileged market estimates, 85% split



Accounting for efficiency wages, adjusted GWGs $\neq$ pooled

Comparing estimates from pooled OLS to endogenous switching regression, 85% split



Summary of the results

What was shown and what was not shown (due to time constraints)

- Results are qualitatively the same with other sample splits (arbitrary thresholds or Cramer split)

Summary of the results

What was shown and what was not shown (due to time constraints)

- Results are qualitatively the same with other sample splits (arbitrary thresholds or Cramer split)
- We test for

Summary of the results

What was shown and what was not shown (due to time constraints)

- Results are qualitatively the same with other sample splits (arbitrary thresholds or Cramer split)
- We test for
  - two regimes (if they exist) → they always do

Summary of the results

What was shown and what was not shown (due to time constraints)

- Results are qualitatively the same with other sample splits (arbitrary thresholds or Cramer split)
- We test for
  - two regimes (if they exist) → they always do
  - significance of gender in the selection equation (joint significance on all interactions) → they always are

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)
  - most countries have higher adjusted GWG in privileged market (consistent with distributional analyses of GWG)

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)
  - most countries have higher adjusted GWG in privileged market (consistent with distributional analyses of GWG)
- Lower estimates adjusted GWG in standard market is a good news: most of the market "discriminates" less

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)
  - most countries have higher adjusted GWG in privileged market (consistent with distributional analyses of GWG)
- Lower estimates adjusted GWG in standard market is a good news: most of the market "discriminates" less

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)
  - most countries have higher adjusted GWG in privileged market (consistent with distributional analyses of GWG)
- Lower estimates adjusted GWG in standard market is a good news: most of the market "discriminates" less $\rightarrow$ policy implications for gender mainstreaming policies

Summary of the results

What was shown and what was not shown (due to time constraints)

- Comparative results – intuitively – make sense:
  - countries with more labor market segmentation have OLS more off (e.g. transition economies, Southern Europe)
  - most countries have higher adjusted GWG in privileged market (consistent with distributional analyses of GWG)
- Lower estimates adjusted GWG in standard market is a good news: most of the market "discriminates" less → policy implications for gender mainstreaming policies
- In some of the markets, virtually all of the "discrimination" is from the gendered labor market segmentation, wages are equal.

Conclusion

Starting point: efficiency wages may interact with other sources of labor market inequality (e.g. biasing estimates of wage gaps). We look at gender (common in all countries, prevalent wage gaps).

Conclusion

Starting point: efficiency wages may interact with other sources of labor market inequality (e.g. biasing estimates of wage gaps). We look at gender (common in all countries, prevalent wage gaps).

We find that:

- estimates which abstract from labor market segmentation bias estimates of GWG;
- access to the privileged market is gendered;
- and that wage inequalities differ across markets.

Conclusion

Starting point: efficiency wages may interact with other sources of labor market inequality (e.g. biasing estimates of wage gaps). We look at gender (common in all countries, prevalent wage gaps).

We find that:

- estimates which abstract from labor market segmentation bias estimates of GWG;
- access to the privileged market is gendered;
- and that wage inequalities differ across markets.

Ahead of us:

- More insights on the properties of this estimator
- Alternative optimization algorithms (FIML? Bayesian?)

Questions?

Thank you for your attention!

w: grape.org.pl
t: grape_org
f: grape.org
e: j.tyrowicz@grape.org.pl (& kbech@sgh.waw.pl)