



# Beyond means: Distributional analysis of gender pay gaps

Valentina Andrade

Advisor: Tomás Rau (PUC)

Co-advisor: Pablo Muñoz (UCH)

**Instituto de Economía - PUC**

December 11, 2023

# Roadmap



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- 1** Motivation
- 2** Model
- 3** Results

- 4** Discussion
- 5** Conclusion
- 6** References

# The gender wage gap increases with quantiles

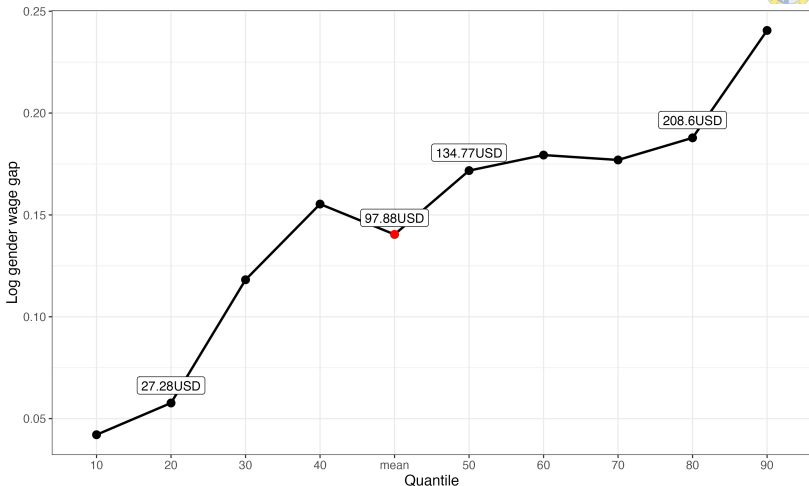


Figure: Distributional Gender Pay Gap in Chile (2013-2018).

Note: Gender pay gap is the difference in log earnings percentiles between males and females.

# Related literature



- Studies with a more “distributional” focus center on examining changes in the top salaries.
- The main explanation has to do with the impact of individuals’ characteristics (Katz & Murphy, 1992).
- What is the role of firms? The gender wage inequalities literature has emphasized this. (Goldin, 2014; Goldin et al., 2017)
  - 1 Two-way FE (Abowd et al., 1999)  $\implies$  Card et al. (2016)
  - 2 Decomposition methods (Blinder, 1973; Oaxaca, 1973)  $\implies$  Blau and Kahn (2017)
- Primarily focus on **means**

# Understanding (gender) wage gap (Card et al., 2016)



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

Card et al., 2016 distinguishes between two approaches to address firm-specific pay policies that can be important for understanding the gender wage gap

## (1) Sorting channel (between)

- 1 Women are less likely to find jobs at higher-paying firms
- 2 Wage gain for given firm-to-firm transition is smallest for women than men

## (2) Bargaining channel (within)

- Wage-setting power of firms and the possibility that women are **offered or negotiate** systematically lower wages at a given firm

# Basic framework



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

## Wage-setting (Abowd et al., 1999; Card et al., 2016)

$$\ln w_{it} = \underbrace{\alpha_j}_{\text{persons effect}} + \underbrace{\psi_{J(i,t)}^{G(i)}}_{\text{gender firm effects}} + \underbrace{X'_{it}\beta^{G(i)}}_{\text{gender returns to the covariates}} + r_{it} \quad (1)$$

- $\psi_{J(i,t)}^{G(i)} \equiv \gamma^{G(i)} \bar{S}_{J(i,t)}$
- $r_{it} = \gamma^{G(i)} (\phi_{J(i,t)} + m_{J(i,t)}) + \varepsilon_{it}$

## Decomposition (Blinder, 1973; Card et al., 2016; Oaxaca, 1973)

$$E[\Psi_{J(i,t)}^M | \text{male}] - E[\Psi_{J(i,t)}^F | \text{female}] \quad (2)$$

# Decomposing the Effect of Firm-Level Wage Premiums



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

$$E[\Psi_{J(i,t)}^M | male] - E[\Psi_{J(i,t)}^F | female] \quad (3)$$

- 1 Difference in mean values of the wage premium **within** the groups
- 2 Group difference in the wage premium (**between**)

# Decomposing the Effect of Firm-Level Wage Premiums



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

$$E[\Psi_{J(i,t)}^M | male] - E[\Psi_{J(i,t)}^F | female] + E[\Psi_{J(i,t)}^F | male] - E[\Psi_{J(i,t)}^F | male]$$

$$E[\Psi_{J(i,t)}^F - \Psi_{J(i,t)}^M | male] + E[\Psi_{J(i,t)}^F | male] - E[\Psi_{J(i,t)}^F | female] \quad (4)$$



# Decomposing the Effect of Firm-Level Wage Premiums



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

$$\underbrace{E[\Psi_{J(i,t)}^F - \Psi_{J(i,t)}^M | male]}_{\text{Bargaining power (within)}} + \underbrace{E[\Psi_{J(i,t)}^F | male] - E[\Psi_{J(i,t)}^F | female]}_{\text{Sorting effect (between)}}$$

- 1 Average bargaining power (*within firms*):** Differential not explained by these difference observed in characteristic of women and men (across the distribution held by men)

# Decomposing the Effect of Firm-Level Wage Premiums



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

$$\underbrace{E[\Psi_{J(i,t)}^F - \Psi_{J(i,t)}^M | male]}_{\text{Bargaining power (within)}} + \underbrace{E[\Psi_{J(i,t)}^F | male] - E[\Psi_{J(i,t)}^F | female]}_{\text{Sorting effect (between)}}$$

- 1 **Average bargaining power** (*within firms*): Differential not explained by these difference observed in characteristic of women and men (across the distribution held by men)
- 2 **Average of sorting channel** (*between firms*): Differences in average wage premium across jobs held by men versus women

# Distributional analysis (Firpo et al., 2009b)



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

When the statistic of interest is a specific real-valued function of the outcome distribution,

$$IF(y; \nu(F_Y)) = \lim_{\varepsilon \rightarrow 0} \frac{[\nu((1 - \varepsilon) \cdot F_Y + \varepsilon \cdot \delta_y) - \nu(F_Y)]}{\varepsilon} = \frac{\partial \nu(\nu(F_Y))}{\partial \varepsilon} \quad (5)$$

With  $0 \leq \varepsilon \leq 1$  and where:

- $F$  signifies the cumulative distribution function of variable  $Y$ .
- $\delta_y$  is a distribution concentrated solely at the value  $y$ .

$$\begin{aligned} RIF(y; \nu(F_Y)) = \\ \nu(F_Y) + \int IF(y; \nu(F_Y)) \cdot dF_Y(y) \\ \nu(F_Y) + IF(y; \nu(F_Y)) \end{aligned} \quad (6)$$

# Distributional analysis (Firpo et al., 2009b)



- The intuition:  $F_Y(y) = \int F_{Y|X}(Y | X = x)dF_X(x)$

## Property 1 - Integral relationship

$$\nu(F_Y) = \int RIF(y; \nu(F_Y))dF_Y(y) = E[RIF(y; \nu(F_Y))]$$

## Theorem 1 - Integration for the Marginal Effect of a change in the distribution

$$\begin{aligned}\nu(F_Y) &= \int E(RIF(y; \nu(F_Y)) | X = x) dF_X(x) \\ &= E\{E(RIF(y; \nu(F_Y)) | X = x)\}\end{aligned}$$

# Extending AKM to the case of quantiles

## ■ AKM version

$$\text{RIF}(y_i; q_\tau) = \alpha_{i\tau} + X_i' \beta_\tau + \Psi_{J(i,t),\tau} + \varepsilon_{i\tau}, \quad E(\varepsilon_{i\tau}) = 0$$



# Extending AKM to the case of quantiles



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- AKM version

$$\text{RIF}(y_i; q_\tau) = \alpha_{i\tau} + X_i' \beta_\tau + \Psi_{J(i,t),\tau} + \varepsilon_{i\tau}, \quad E(\varepsilon_{i\tau}) = 0$$

- Property 1

$$q_\tau = E(\text{RIF}(y_i; q_\tau)) = E(\alpha_{i\tau}) + E(X_i' \beta_\tau) + E(\Psi_{J(i,t),\tau}) + E(\varepsilon_{i\tau})$$

# Extending AKM to the case of quantiles



## ■ AKM version

$$\text{RIF}(y_i; q_\tau) = \alpha_{i\tau} + X_i' \beta_\tau + \Psi_{J(i,t),\tau} + \varepsilon_{i\tau}, \quad E(\varepsilon_{i\tau}) = 0$$

## ■ Property 1

$$q_\tau = E(\text{RIF}(y_i; q_\tau)) = E(\alpha_{i\tau}) + E(X_i' \beta_\tau) + E(\Psi_{J(i,t),\tau}) + E(\varepsilon_{i\tau})$$

## ■ Theorem 1 (LIE)

$$\begin{aligned} q_\tau &= E\{E(\text{RIF}(y_i; q_\tau) \mid \cdot)\} \\ &= E\{E(\alpha_{i\tau}) + E(X_i' \beta_\tau) + E(\Psi_{J(i,t),\tau}) + E(\varepsilon_{i\tau}) \mid \cdot\} \\ &= E\{E(\alpha_{i\tau}) \mid \cdot\} + E\{E(X_i' \beta_\tau) \mid \cdot\} + E\{E(\Psi_{J(i,t),\tau}) \mid \cdot\} \end{aligned}$$

With  $E(\varepsilon_{i\tau}) = 0$  and  $\cdot = \alpha_{i\tau}, X_i' \beta_\tau, \Psi_{J(i,t),\tau}$

# Extending CCK Descomposition



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

In general

$$\begin{aligned} \Delta_o^\nu &= \nu(F_{Y_1|D_1}) - \nu(F_{Y_0|D_0}) \\ &= \underbrace{\nu(F_{Y_1|D_1}) - \nu(F_{Y_1|D_0})}_{\Delta_S^\nu} + \underbrace{\nu(F_{Y_1|D_0}) - \nu(F_{Y_0|D_0})}_{\Delta_X^\nu} \end{aligned}$$

For firm premiums

$$\Delta_\Psi^\nu = \underbrace{\mathbb{E}[\Psi | D = 1]}'(\gamma_1^\nu - \gamma_0^\nu)}_{\Delta_S^\nu} + \underbrace{(\mathbb{E}[\Psi | D = 1] - \mathbb{E}[\Psi | D = 0])}'\gamma_0^\nu}_{\Delta_X^\nu} \quad (7)$$



# 1. Data



- Matched employer-employee data from chilean Unemployment Insurance (UI) Registry
- 42,796,782 observations, representing a panel of individuals per month from January 2013 to December 2018 ( $t = 60$ )
- Monthly information of taxable wages
- Data solely covers the formal sector covered by labor code
- Limit sample to individuals whose earnings meet or exceed the minimum wage requirement for each respective period

## 2. Goodness of fit



Table: Summary of Estimated TWFE Models by method for all sample, females, and males (2013-2018)

	Indicator	Mean	10	20	50	80	90
all	N	42,786,782	42,786,782	42,786,782	42,786,782	42,786,782	42,786,782
	Person FE ( $\theta$ )	1,198,798	1,198,798	1,198,798	1,198,798	1,198,798	1,198,798
	Firm FE ( $\phi$ )	80,417	80,417	80,417	80,417	80,417	80,417
	$R^2$ adjusted	0.88	0.49	0.60	0.72	0.76	0.77
	RMSE	0.22	0.34	0.39	0.46	0.65	0.85
F	N	16,542,868	16,542,868	16,542,868	16,542,868	16,542,868	16,542,868
	Person FE ( $\theta$ )	486,794	486,794	486,794	486,794	486,794	486,794
	Firm FE ( $\phi$ )	80,417	80,417	80,417	80,417	80,417	80,417
	$R^2$ adjusted	0.89	0.50	0.59	0.72	0.77	0.76
	RMSE	0.21	0.31	0.29	0.42	0.59	0.83
M	N	26,237,430	26,237,430	26,237,430	26,237,430	26,237,430	26,237,430
	Person FE ( $\theta$ )	711,681	711,681	711,681	711,681	711,681	711,681
	Firm FE ( $\phi$ )	80,417	80,417	80,417	80,417	80,417	80,417
	$R^2$ adjusted	0.88	0.50	0.61	0.71	0.76	0.77
	RMSE	0.23	0.34	0.44	0.46	0.66	0.79

### 3. OB decomposition



Table: Oaxaca-Blinder decompositions for Firm FE (2013-2018)

Method	Gender Pay gap	Male premium	Female premium	Gender premium gap	Sorting	Bargaining
10	0.091	0.125	0.177	-0.052	0.0260	-0.0780
20	0.151	0.178	0.199	-0.021	0.0485	-0.0700
50	0.241	0.244	0.144	0.100	0.0720	0.0280
80	0.276	0.206	0.114	0.092	0.0575	0.0345
90	0.314	0.242	0.098	0.144	0.0610	0.0825

Note: sorting and bargaining are calculated using both counterfactuals

#### (1) What is the role of firms?

Firm premium benefits women for lower quantiles, while for higher quantiles and above the 50th percentile, it benefits men

### 3. OB decomposition



Table: Oaxaca-Blinder decompositions for Firm FE (2013-2018)

Method	Gender Pay gap	Male premium	Female premium	Gender premium gap	Sorting	Bargaining
10	0.091	0.125	0.177	-0.052	0.0260	-0.0780
20	0.151	0.178	0.199	-0.021	0.0485	-0.0700
50	0.241	0.244	0.144	0.100	0.0720	0.0280
80	0.276	0.206	0.114	0.092	0.0575	0.0345
90	0.314	0.242	0.098	0.144	0.0610	0.0825

Note: sorting and bargaining are calculated using both counterfactuals

#### (2) Lower quantiles

Gender pay gap is reduced due to bargaining. Nevertheless, sorting counteracts this effect, benefiting men

### 3. OB decomposition



Table: Oaxaca-Blinder decompositions for Firm FE (2013-2018)

Method	Gender Pay gap	Male premium	Female premium	Gender premium gap	Sorting	Bargaining
10	0.091	0.125	0.177	-0.052	0.0260	-0.0780
20	0.151	0.178	0.199	-0.021	0.0485	-0.0700
50	0.241	0.244	0.144	0.100	0.0720	0.0280
80	0.276	0.206	0.114	0.092	0.0575	0.0345
90	0.314	0.242	0.098	0.144	0.0610	0.0825

Note: sorting and bargaining are calculated using both counterfactuals

#### (3) Upper quantiles and median

Sorting becomes more important than bargaining for explaining gender pay gaps. Both components benefit men in terms of wages



# 4. Unionization?

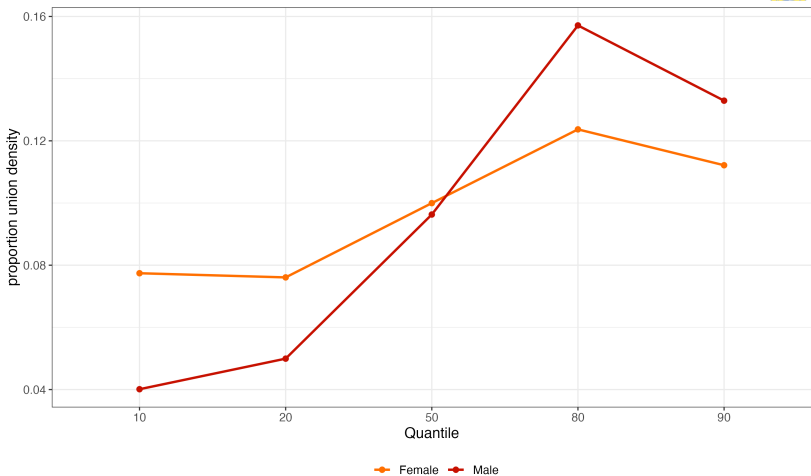


Figure: Union density by sex and quantiles of labor income.  
 Source: CASEN (2017) using private sector employees who meet the criteria of earning at least the minimum wage or more

## 5. Gender-differentiation in labor market

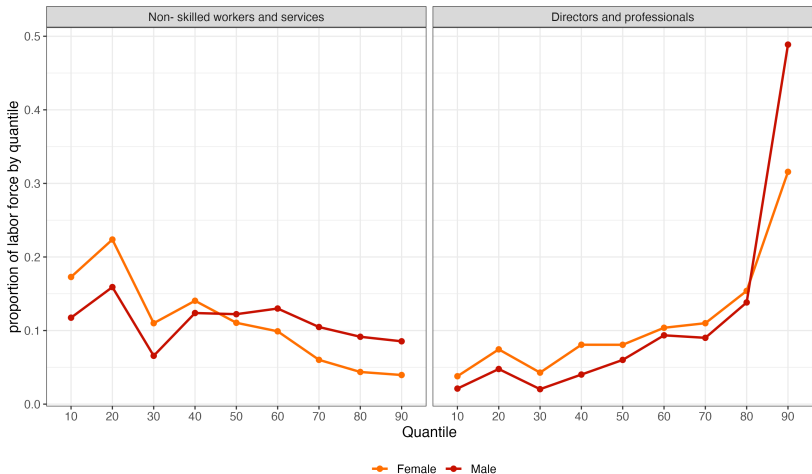


Figure: Labour force participation by sex and quantile.

Source: CASEN (2017) using private sector employees who meet the criteria of earning at least the minimum wage or more

# Conclusion



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- 1 Concerning the gender pay gap, we found that this gap widens as we move up the income quantiles.



# Conclusion



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- 1 Concerning the gender pay gap, we found that this gap widens as we move up the income quantiles.
- 2 In the lowest quantile, firms helps reduce the gap, and in the highest quantile, firms contributes to its increase.

# Conclusion



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- 1 Concerning the gender pay gap, we found that this gap widens as we move up the income quantiles.
- 2 In the lowest quantile, firms helps reduce the gap, and in the highest quantile, firms contributes to its increase.
- 3 We discovered that the bargaining dimension contributes to decrease the firm-driven gender pay gaps at the lower end of the income distribution, but not at the upper end.

# Conclusion



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS

- 1 Concerning the gender pay gap, we found that this gap widens as we move up the income quantiles.
- 2 In the lowest quantile, firms helps reduce the gap, and in the highest quantile, firms contributes to its increase.
- 3 We discovered that the bargaining dimension contributes to decrease the firm-driven gender pay gaps at the lower end of the income distribution, but not at the upper end.
- 4 Next steps: Residual part?  $\implies$  Bargaining (Card et al., 2016) but we must to confirm unionization hypothesis.



# Thanks!

## Beyond means: Distributional analysis of gender pay gaps

Valentina Andrade [vandrade@uc.cl](mailto:vandrade@uc.cl)

Advisor: Tomás Rau (PUC)

Co-advisor: Pablo Muñoz (UCH)



Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms [Publisher: Wiley Online Library]. **Econometrica**, **67**(2), 251–333.



Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations [Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2425]. **Journal of economic literature**, **55**(3), 789–865.



Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates [Publisher: JSTOR]. **Journal of Human resources**, 436–455.



Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., & Setzler, B. (2020). How much should we trust estimates of firm effects and worker sorting?



Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., & Setzler, B. (2023). How much should we trust estimates of firm effects and worker sorting? [Publisher: The University of Chicago Press Chicago, IL]. **Journal of Labor Economics**, **41**(2), 000–000.



Bonhomme, S., Lamadon, T., & Manresa, E. (2019). A distributional framework for matched employer employee data [Publisher: Wiley Online Library]. **Econometrica**, **87**(3), 699–739.



INSTITUTO DE ECONOMÍA  
FACULTAD DE CIENCIAS ECONÓMICAS  
Y ADMINISTRATIVAS



Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018). Firms and labor market inequality: Evidence and some theory [Publisher: University of Chicago Press Chicago, IL]. **Journal of Labor Economics**, **36**, S13–S70.



Card, D., Cardoso, A. R., & Kline, P. (2013). Bargaining and the gender wage gap: A direct assessment [Publisher: IZA Discussion Paper].



Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women [Publisher: Oxford University Press]. **The Quarterly journal of economics**, **131**(2), 633–686.



Firpo, S., Fortin, N. M., & Lemieux, T. (2009a). Supplement to 'unconditional quantile regressions'. **Econometrica Supplemental Material**, **77**.



Firpo, S., Fortin, N. M., & Lemieux, T. (2009b). Unconditional quantile regressions [Publisher: Wiley Online Library]. **Econometrica**, **77**(3), 953–973.



Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition methods in economics. In **Handbook of labor economics** (pp. 1–102, Vol. 4). Elsevier.



Goldin, C. (2014). A grand gender convergence: Its last chapter [Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203]. **American economic review**, **104**(4), 1091–1119. Retrieved November 3, 2023, from <https://www.aeaweb.org/articles?id=10.1257/aer.104.4.1091>



Goldin, C., Kerr, S. P., Olivetti, C., & Barth, E. (2017). The expanding gender earnings gap: Evidence from the LEHD-2000 census [Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203]. **American Economic Review**, **107**(5), 110–114. Retrieved November 3, 2023, from <https://www.aeaweb.org/articles?id=10.1257/aer.p20171065>



Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors [Publisher: MIT Press]. **The quarterly journal of economics**, **107**(1), 35–78.



Oaxaca, R. (1973). Male-female wage differentials in urban labor markets [Publisher: JSTOR]. **International economic review**, 693–709.

# Identification strategy (Card et al., 2018)



- Identifying age and time effects: (1) difficult reconstructing a worker's employment history and (2) not clear employment gaps are exogenous.
- Normalization: choice of normalization is important for values of person-effects and time-varying controls.
- **Limited Mobility Bias:** (Bonhomme et al., 2019, 2020, 2023) parameters that are solely identified by workers who move across firms.
  - FE estimates of the **contribution of firm effects** to wage inequality are **upward biased**.
  - FE estimates of **contribution of the sorting of workers to firms** are **biased downward**.
- **Exogeneous Mobility:** worker and firm effect will be biased unless worker mobility is orthogonal to time-varying residual components of wages.
- Additive separability: it assumption about the proportional firm wage effect for all workers. (Card et al., 2013)
- **Distributions?**



# Distributional Analysis (Firpo et al., 2009a)



## Property 1

$$\begin{aligned}\nu(F_Y) &= \int RIF(y; \nu, F_Y) dF_Y(y) \\ &= \iint RIF(Y; \nu, F_Y) \cdot dF_{Y|X}(y | X = x) \cdot dF_X(x) \\ &= \int E[RIF(Y; \nu, F_Y) | X = x] \cdot dF_X(x)\end{aligned}$$

# Distributional Analysis (Fortin et al., 2011)



## Equation 7

Let  $Y = f(\Phi, \varepsilon) = \alpha_i + \Psi_{J(i,t)}^{D_i} + X' \beta^{D_i} + \varepsilon$

$$\begin{aligned}
 \Delta'_O &= \mathbb{E}[Y \mid D = 1] - \mathbb{E}[Y \mid D = 0] \\
 &= \mathbb{E}[\mathbb{E}(Y \mid \Phi, D = 1) \mid D = 1] - \mathbb{E}[\mathbb{E}(Y \mid \Phi, D = 0) \mid D = 0] \\
 &= (\mathbb{E}[\Phi \mid D = 1]' \gamma'_1 + \mathbb{E}[\varepsilon_1 \mid D = 1]) - (\mathbb{E}[\Phi \mid D = 0]' \gamma'_0 + \mathbb{E}[\varepsilon_0 \mid D = 0]),
 \end{aligned} \tag{8}$$

where  $\mathbb{E}[\varepsilon_s \mid D = s] = 0$  because  $\mathbb{E}[\varepsilon_s \mid \Phi, D = s] = 0$ , so the expression reduces to  $\Delta'_O = \mathbb{E}[\Phi \mid D = 1]' \gamma'_1 - \mathbb{E}[\Phi \mid D = 0]' \gamma'_0$ . Thus, by adding and subtracting  $\mathbb{E}[\Phi \mid D = 1]' \gamma'_0$  we get

$$\Delta'_O = \underbrace{\mathbb{E}[\Phi \mid D = 1]' (\gamma'_1 - \gamma'_0)}_{\Delta'_{S,OB}} + \underbrace{(\mathbb{E}[\Phi \mid D = 1] - \mathbb{E}[\Phi \mid D = 0])' \gamma'_0}_{\Delta'_{X,OB}}.$$