Inequality of Opportunity and the Probability of Being Very Rich or Very Poor

Alessio Rebechi [1], Nicholas Rohde [1], Gordon Anderson [2]

[1]Department of Accounting, Finance and Economics, Griffith University

[2]Department of Economics, University of Toronto

III/LIS Conference
Session "Inequality of Opportunity"
24th February 2023

Outline

Beyond the Mean

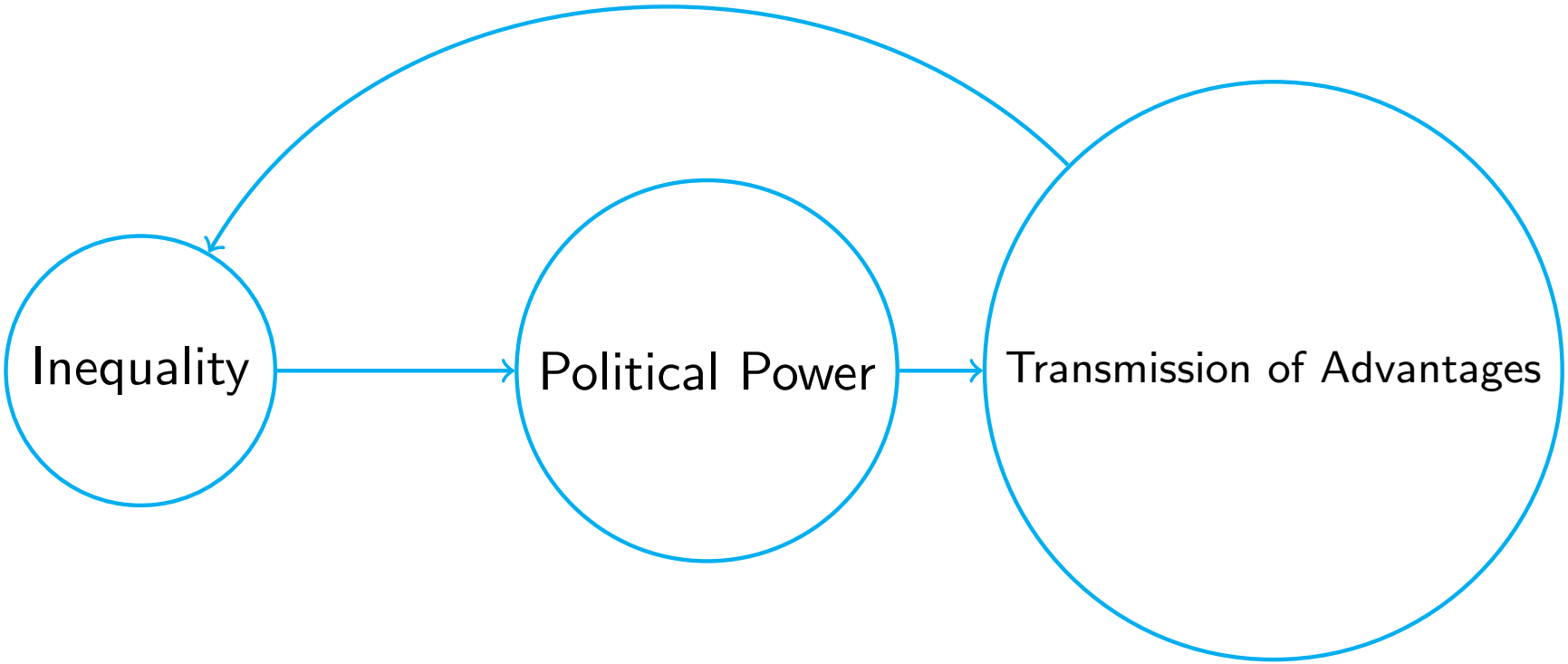
- We study the link between background characteristics and the tails of income distribution using some full-distributional regression models.
- Most of the IOp empirical research has focused on conditional means, neglecting the relative frequency of extreme values.
- We argue that more attention should be given to the composition of the left and right tails since they disproportionately influence public perceptions of economic disparity.
- We look at other features of the distribution (e.g., the variance) to capture additional information that the traditional expected values-based approach would have missed.
- A complementary approach to reconcile traditional methods of measuring IOp with general perceptions of inequality as unequal chances is proposed.

Inequality of Opportunity and Distributional Tails

- IOp is defined as the inequality due to variables beyond individual control such as gender, race or parental socio-economic status (Ferreira & Gignoux, 2011; Roemer, 1998; Roemer & Trannoy, 2016, among others).
- The relevance of extreme values and heavy tails has been discussed in the analysis of the income distribution (Bossert *et al.*, 2021; Ibragimov & Ibragimov, 2018; Schluter, 2012), however, this framework has not been extended to the IOp literature.
- Emphasizing the tails and their composition is particularly important in providing a more realistic picture of IOp and its implications.
- For example, the over-representation of minorities among poor people is related to disparities in crimes statistics that are not immediately evident from mean based statistics.

Why the Tails Matter?

Inequality & Politics (Milanovic, 2019)



- Inequality & Politics literature (Gethin *et al.*, 2021; Hacker & Pierson, 2010; Milanovic, 2019; Piketty, 2017, 2020; Piketty & Saez, 2006).

Why the Tails Matter?

- Economic analysis of populism: persistence of group-based inequality increases anxiety, distrust $\Rightarrow$ populist support (Bossert *et al.*, 2019; Guiso *et al.*, 2017; Guriev & Papaioannou, 2020; Rodrik, 2021).
- Perception of inequality and fairness $\Rightarrow$ individual attitudes on redistributive policies (Alesina & Angeletos, 2005; Alesina & La Ferrara, 2005; Benabou & Ok, 1998; Benabou & Tirole, 2005; Piketty, 1995).

Outline

Household, Income and Labor Dynamics in Australia (HILDA)

- Last release of HILDA data (2021). HILDA is a panel study that started in 2001, it collects information on different aspects of life from more than 17,000 Australians each year.
- Key variables:
  - (log) household income, which is defined after governmental taxes and transfers, corrected for age and inflation, and standardized using the square-root adult equivalence scale;
  - (log) average weekly wage and salary income (imputed) from all forms of paid employment over the time, defined before taxation and governmental transfers and corrected for age and inflation;
- Circumstances: gender, immigration status, parental SES and family environment.
- Sample: Household income sample 242,900 observations; Weekly earnings sample more than 120,000 observations. The observations are taken over a period of 20 years from 2001 to 2020.

Outline

Full-Distributional Approach

- Log-Normal distribution PDF:

$$f(y) = \frac{1}{(y\sigma\sqrt{(2\pi)})} \exp(-\frac{(\ln(y) - \mu)^2}{2\sigma^2}) \tag{1}$$

- Heteroskedastic linear regression (similar to Jenkins (2007)):

$$\hat{\mu}_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \tag{2}$$

$$\hat{\sigma_i^2} = \exp\left(\hat{\theta}_0 + \sum_{l=1}^{m} \hat{\theta}_l x_{il}\right) \tag{3}$$

- From the estimated parameters, we calculate the differentials between the probability of being in the tails of the unconditional distribution and the probability of being in the tails of the conditional distribution.

Full-Distributional Approach

- Cut-offs: poverty line (z) as 50% of median income/wage and top 1% of the income wage distribution (k), parametrically estimated.
- Z-scores:

$$z_{poor} = \frac{z - \hat{\mu}_{x_i}}{\sqrt{\sigma^2_{x_i}}} \tag{4}$$

$$z_{rich} = \frac{k - \hat{\mu}_{x_i}}{\sqrt{\sigma^2_{x_i}}} \tag{5}$$

- Probability of being in the tails:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt[2]{2\pi}} e^{x^2/2} dx \tag{6}$$

$$Pr(y > k) = 1 - \phi(z_{rich}) \tag{7}$$

$$Pr(y < z) = \phi(z_{poor}) \tag{8}$$

Outline

Heteroskedastic Regressions Results

	(1)		(2)	
	Log Household Income		Log Weekly Wage	
	$\hat{\mu}$	$\ln(\sigma^2)$	$\hat{\mu}$	$\ln(\sigma^2)$
Female	-0.0501***	-0.00346	-0.415***	0.182***
	(0.00237)	(0.0134)	(0.00394)	(0.0119)
Refugee	-0.0780***	0.0908	-0.0149	0.169**
	(0.0107)	(0.0572)	(0.0180)	(0.0582)
Indigenous origin	-0.255***	-0.105	-0.0233	-0.0602
	(0.0108)	(0.0593)	(0.0211)	(0.0732)
Immigrant	0.0315***	0.170***	0.0800***	-0.0980***
	(0.00510)	(0.0276)	(0.00795)	(0.0242)
Mother immigrant	-0.00491	-0.0421*	-0.00962	-0.0336
	(0.00394)	(0.0214)	(0.00657)	(0.0195)
Father immigrant	0.000132	-0.0121	-0.0137*	0.0107
	(0.00367)	(0.0192)	(0.00614)	(0.0183)
First language learned: English	0.137***	-0.0163	0.101***	0.0110
	(0.00576)	(0.0297)	(0.00870)	(0.0283)
Parents divorced/separated	-0.0797***	0.00790	-0.0300***	0.000133
	(0.00380)	(0.0211)	(0.00609)	(0.0193)
Oldest child	0.0323***	0.0114	0.0447***	0.0402**
	(0.00251)	(0.0139)	(0.00414)	(0.0125)
Non-biological father	-0.136***	-0.0126	-0.178***	0.0351
	(0.00848)	(0.0539)	(0.0153)	(0.0553)
Non-biological mother	0.0179	0.0543	0.0776***	-0.0770
	(0.0107)	(0.0631)	(0.0182)	(0.0571)
Father university	0.159***	0.0821***	0.0680***	0.0879***
	(0.00363)	(0.0204)	(0.00610)	(0.0167)
Mother university	0.0841***	-0.0695**	0.00930	0.171***
	(0.00381)	(0.0219)	(0.00710)	(0.0184)
Father employed	0.186***	0.0421	0.0262**	-0.0649*
	(0.00521)	(0.0271)	(0.00930)	(0.0261)
Mother employed	0.0756***	-0.197***	0.0403***	-0.0566***
	(0.00247)	(0.0137)	(0.00407)	(0.0124)
Constant	10.70***	-0.936***	7.183***	-0.893***
	(0.00960)	(0.0513)	(0.0150)	(0.0462)
Observations	242,994		129,651	
$\chi 2$ for mean model test	41766.0		20345.5	
$\chi 2$ for heteroskedasticity test	619.2		635.8	
p-value for heteroskedasticity test	0.0000		0.0000	

Gender

Table: Conditional Probabilities of being in the Tails of the distribution by Gender

	Probability of being poor	Probability of being in the top 1%
	Log Household Income	
Female	0.126	0.005
Male	0.110	0.007
Ratio Female/Male	1.150	0.775
T-value Female/Male	12.621	4.841
P-value	0.000	0.000
	Log Weekly Wage	
Female	0.257	0.004
Male	0.092	0.010
Ratio Female/Male	2.795	0.363
T-value Female/Male	80.031	14.148
P-value	0.000	0.000

Notes: The table presents the conditional probabilities of being in the top 1% or being under the poverty line by gender for the log household income and the log weekly wage. We also report the ratios, the t-statistics, and the p-values for the two sample.

Immigration Status

Table: Conditional Probabilities of Being in the Tails of the Distribution by Immigration Status

	Probability of being poor	Probability of being top 1%
	Log Household Income	
Immigrant	0.126	0.010
Non-Immigrant	0.116	0.005
Ratio Immigrant/Non-Immigrant	1.081	2.051
T-value Immigrant/Non-Immigrant	5.671	10.996
P-value	0.000	0.000
	Log Weekly Wage	
Immigrant	0.132	0.006
Non-Immigrant	0.171	0.006
Ratio Immigrant/Non-Immigrant	0.773	0.978
T-value Immigrant/Non-Immigrant	15.986	0.253
P-value	0.000	0.800

Notes: The table presents the conditional probabilities of being in the top 1% or being under the poverty line by immigration status for the log of household income and the log of weekly wage. We also report the ratios, the t-statistics, and the p-values for the two sample.

Parental SES

Table: Conditional Probabilities of Being in the Tails of the Distribution by Father's Education

	Probability of being poor	Probability of being top 1%
	Log Household Income	
Father university	0.086	0.014
Father without university degree	0.125	0.005
Ratio Father university/non-university	0.689	2.730
T-value Father university/ non-university	24.128	14.086
P-value	0.000	0.000
	Log Weekly Wage	
Father university	0.153	0.010
Father without university degree	0.166	0.006
Ratio Father university/non-university	0.923	1.732
T-value Father university/ non-university	4.753	5.803
P-value	0.000	0.000

Notes: The table presents the conditional probabilities of being in the top 1% or being under the poverty line by father's education for the two periods considered. We also report the ratios, the t-statistics, and the p-values for the two sample.

Implications

- Over-representation of men in top managerial positions and decision-making roles contributes to maintain their positional advantages with respect to women.
- Immigrants more likely to be in the top percentile. This can contribute to both stereotypes ('all doctors are Indian') and feeling of being 'out of place', resentment for being overtaken by minorities.
- Over-representation of minorities and women in the left tails shapes individual perceptions (of themselves and others) and beliefs around these groups (statistical discrimination) consistently excluding them from productive opportunities.
- The transmission of educational and financial advantage across generations further reinforce the cycle of privilege.
- Democratic functioning: the composition of the top tail affect public policies, with technocratic governments having policy preferences that often don't align with the general needs of the population.

Outline

Summary

- The mean based IOp approach provides a limited view of the effects of circumstances compared to the full distributional approach.
- We model outcome variance as a function of circumstances, which allows us to detect heterogeneity among individuals from the same type, particularly important to capture the effect on the extremes.
- The utility of our approach in providing a more comprehensive picture of IOp is confirmed by our results, especially those regarding immigration status and parental background.
- Political salience of Inequality of Oppportunity.
- More comprehensive picture of the distributional impact of public policies: capture heterogeneous effects and properly identify those who gain and those who lose from policies implementation (Carneiro *et al.*, 2003; Heckman, 2001).

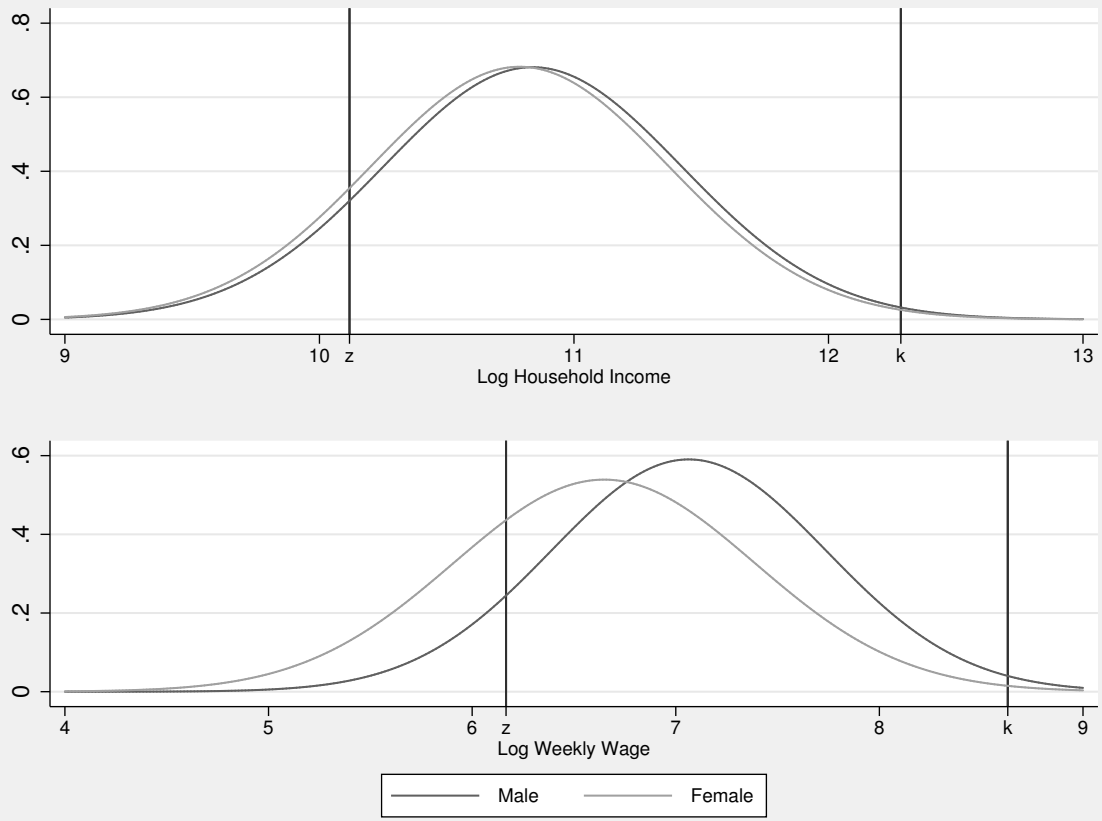
Acknowledgment

Thanks for your attention.

Outline

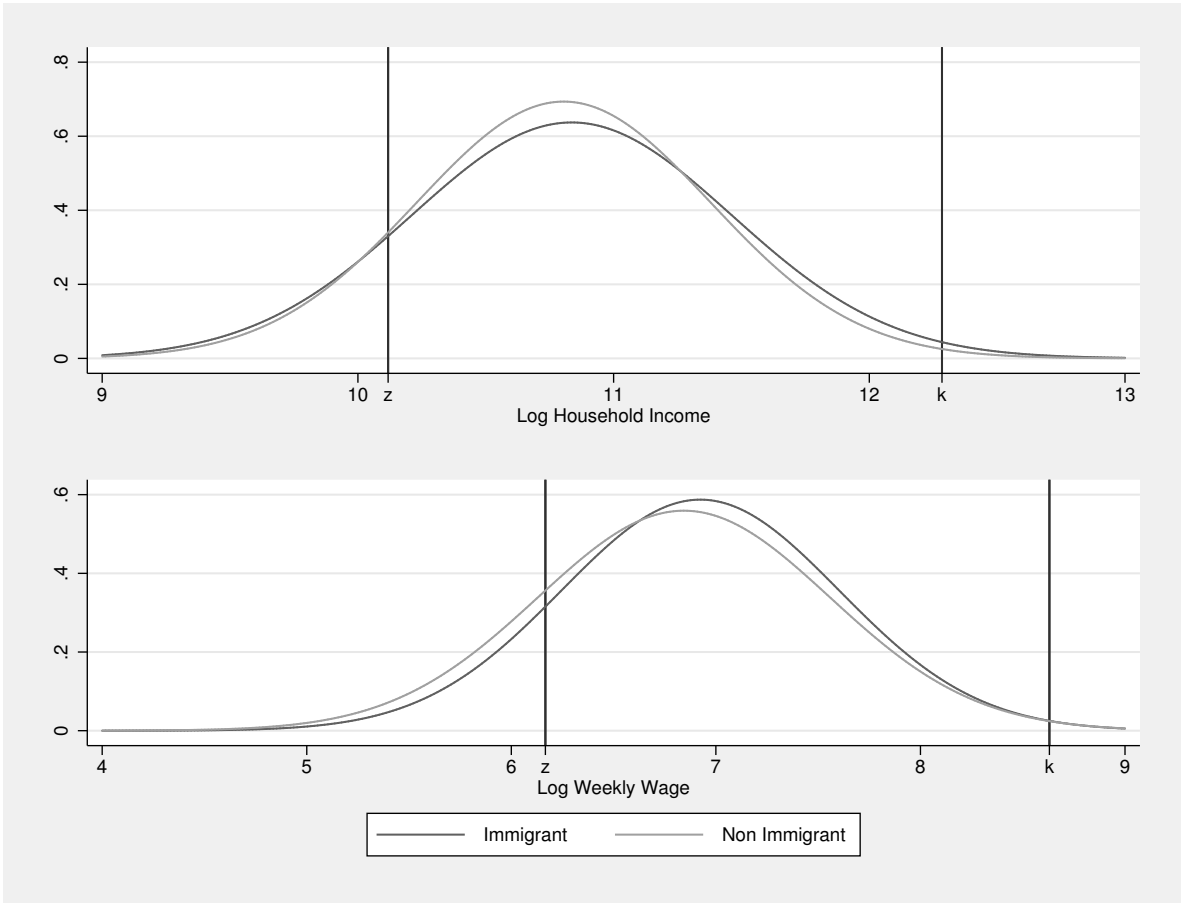
Gender

Figure: Conditional Density Functions by Gender



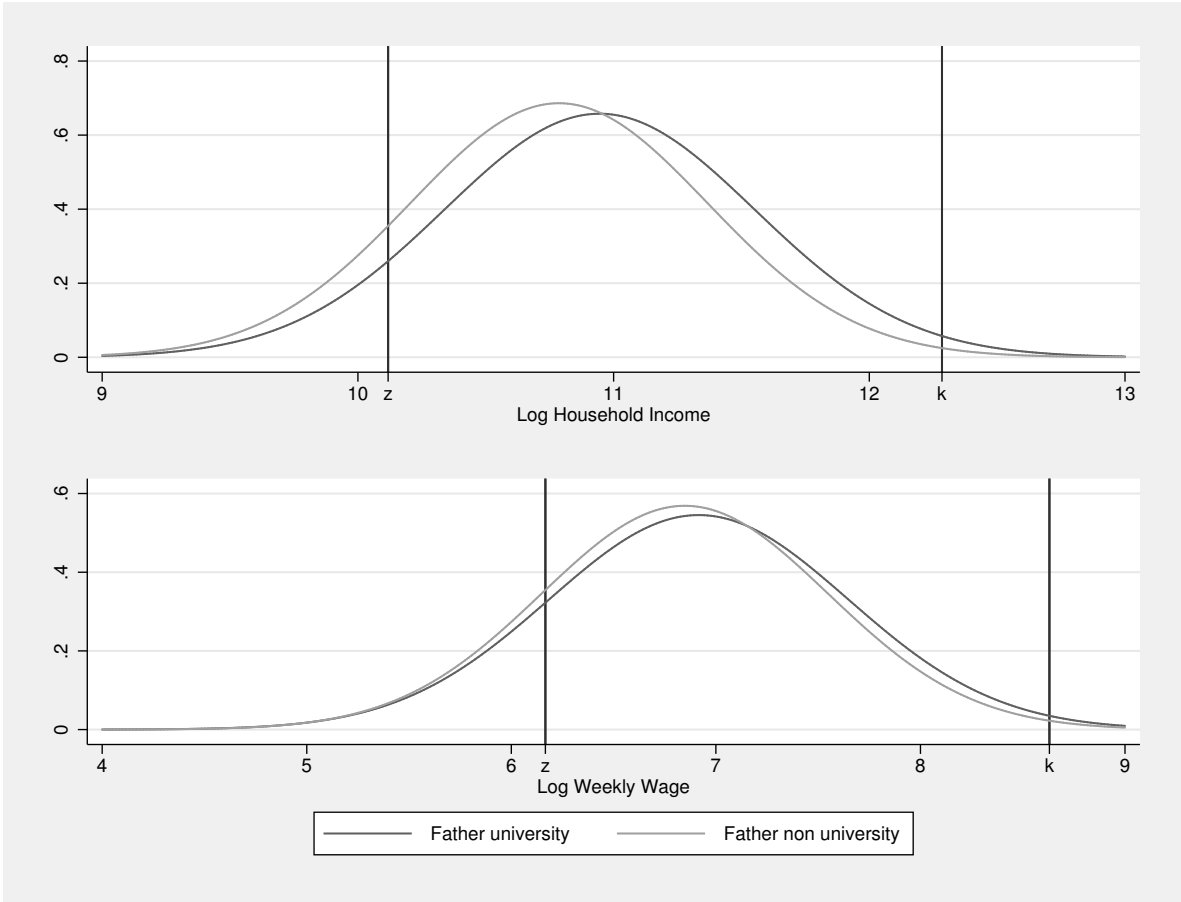
Immigration Status

Figure: Conditional Density Functions by Immigration Status



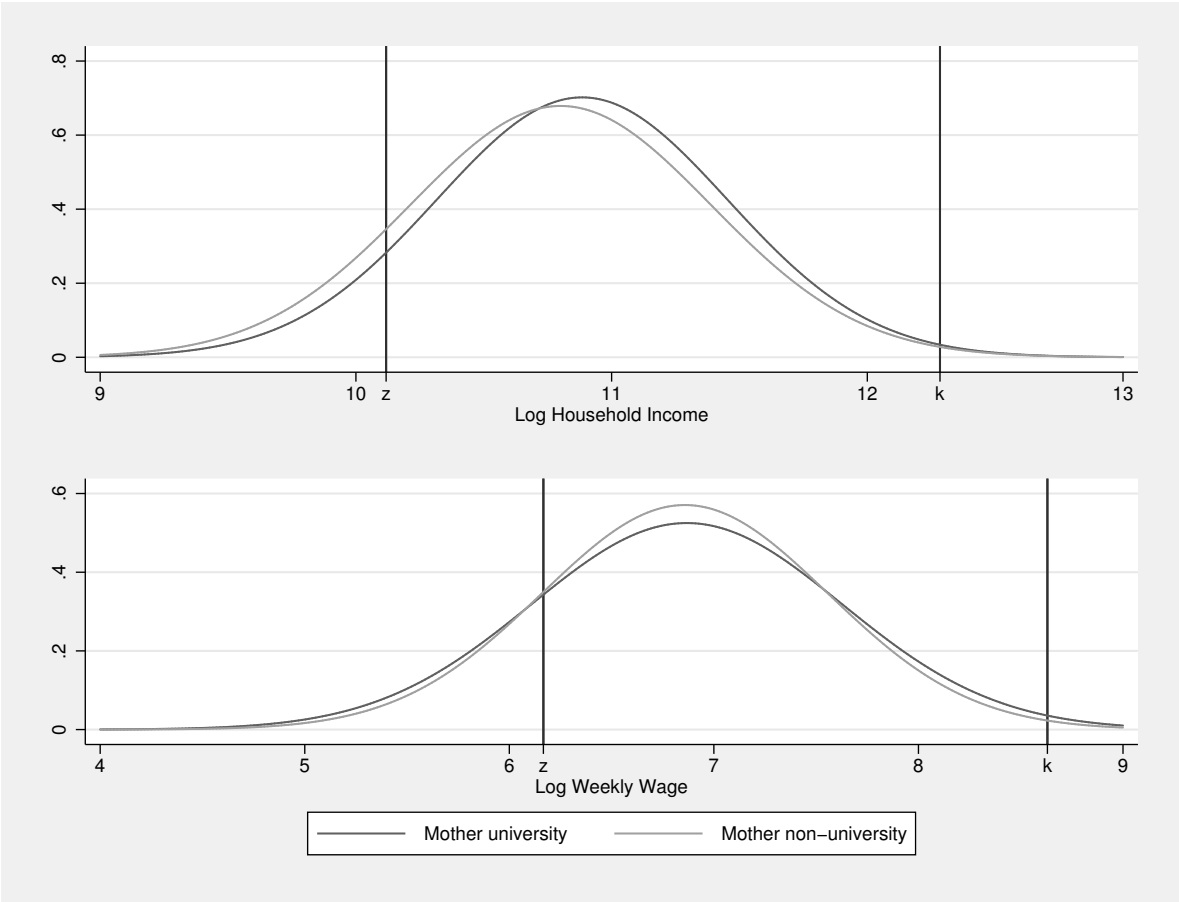
Parental SES

Figure: Log Household Income Conditional Density Functions by Father's Education



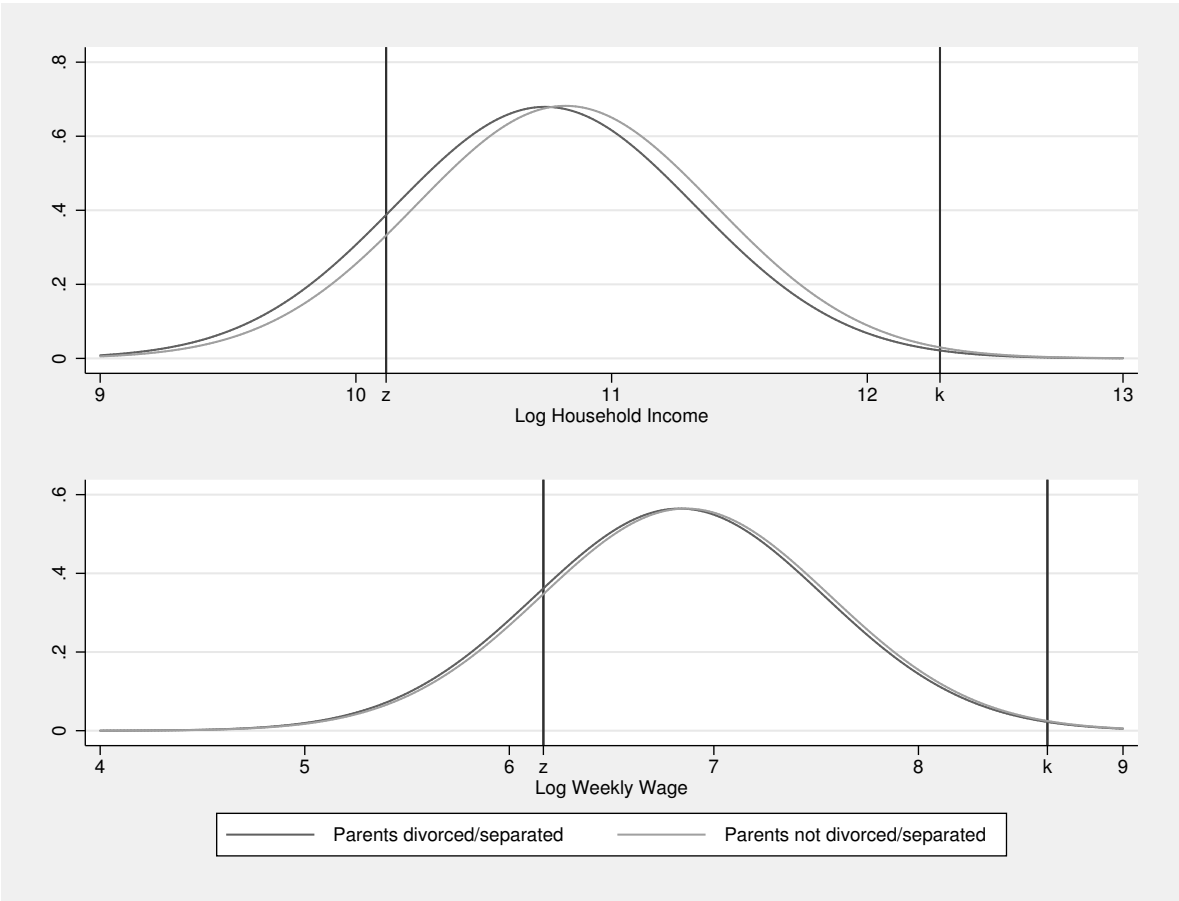
Parental SES

Figure: Log Household Income Conditional Density Functions by Mother's Education



Family Environment

Figure: Conditional Density Functions by Parents' Marital Status



Homoskedastic Linear Regression Results

	(1)	(2)
	Log Household Income	Log Weekly Wage
Female	-0.0515***	-0.415***
	(0.00239)	(0.00396)
Refugee	-0.0804***	-0.0184
	(0.0108)	(0.0183)
Indigenous origin	-0.268***	-0.0214
	(0.0109)	(0.0210)
Immigrant	0.0312***	0.0793***
	(0.00513)	(0.00801)
Mother immigrant	-0.00724	-0.00707
	(0.00392)	(0.00667)
Father immigrant	0.000197	-0.0156*
	(0.00364)	(0.00628)
First language learned: English	0.137***	0.0937***
	(0.00573)	(0.00877)
Parents divorced/separeted	-0.0774***	-0.0258***
	(0.00382)	(0.00610)
Oldest child	0.0335***	0.0453***
	(0.00254)	(0.00418)
Non-biological father	-0.140***	-0.173***
	(0.00898)	(0.0158)
Non-biological mother	0.0259*	0.0646***
	(0.0114)	(0.0187)
Father university	0.159***	0.0689***
	(0.00366)	(0.00599)
Mother university	0.0806***	0.0117
	(0.00387)	(0.00693)
Father employed	0.188***	0.0166
	(0.00524)	(0.00935)
Mother employed	0.0752***	0.0405***
	(0.00247)	(0.00408)
Constant	10.700***	7.204***
	(0.00957)	(0.01511)
Log sigma2		
Constant	-1.060***	-0.684***
	(0.00687)	(0.00594)
Observations	242,994	129,651
$\chi 2$ for mean model test	40032.2	18865.7

Unconditional & Conditional Probabilities

Table: Conditional Probabilities of being in the Tails of the distribution

	Probability of being poor	Probability of being in the top 1%	Probability of being poor	Probability of being in the top 1%
	Log Household Income		Log Weekly Wage	
Female	0.126	0.005	0.257	0.004
Male	0.110	0.007	0.092	0.010
Ratio Female/Male	1.150	0.775	2.795	0.363
T-value Female/Male	12.621	4.841	80.031	14.148
P-value	0.000	0.000	0.000	0.000
Immigrant	0.126	0.010	0.132	0.006
Non-Immigrant	0.116	0.005	0.171	0.006
Ratio Immigrant/Non-Immigrant	1.081	2.051	0.773	0.978
T-value Immigrant/Non-Immigrant	5.671	10.996	15.986	0.253
P-value	0.000	0.000	0.000	0.800
Father university	0.086	0.014	0.153	0.010
Father without university degree	0.125	0.005	0.166	0.006
Ratio Father university/non-university	0.689	2.730	0.923	1.732
T-value Father university/ non-university	24.128	14.086	4.753	5.803
P-value	0.000	0.000	0.000	0.000
Mother university	0.089	0.007	0.178	0.010
Mother without university degree	0.123	0.006	0.161	0.006
Ratio mother university/non-university	0.722	1.190	1.105	1.812
T value Mother university/non-university	19.237	2.204	5.347	5.614
P-value	0.000	0.000	0.000	0.000
Father employed	0.115	0.006	0.162	0.006
Father not employed	0.183	0.002	0.180	0.007
Ratio Father employed/not employed	0.627	3.086	0.905	0.883
T value Father employed/not employed	20.277	10.130	3.550	0.773
P-value	0.000	0.000	0.000	0.440
Mother employed	0.096	0.005	0.155	0.006
Mother not employed	0.145	0.007	0.176	0.006
Ratio Mother employed/not employed	0.664	0.715	0.881	0.961
T value Mother employed/not employed	36.619	6.345	9.988	0.549
P-value	0.000	0.000	0.000	0.583
Parents divorced/separated	0.145	0.004	0.173	0.005
Parents non-divorced/separated	0.115	0.115	0.162	0.006
Ratio Parents divorced/Non-divorced	1.257	0.696	1.066	0.887
T value Parents divorced/Non-divorced	13.154	4.336	3.301	1.100
P-value	0.000	0.000	0.001	0.271