Field Specializations Among Beginning Economists: Are There Gender Differences?

Eva Sierminska and Ronald L. Oaxaca

LISER, DIW, GLO, IZA, U of A; University of Arizona, GLO, IZA, LISER

ASSA Virtual Annual Meeting
January 3-5, 2021

Introduction

- One of the most frequently studied topic on gender disparities is the gender wage gap.
- Likely in second place are gender differences in occupational distributions.
- Along these lines is a major policy interest in gender differences in STEM field representation.
- Because the field of economics embraces the full spectrum from formal abstract reasoning to empirical approaches, documenting and understanding gender differences in economics fields of specialization has the potential to shed light on the more general issue of women in the STEM fields.
- Moreover, there is clearly a more parochial interest among economists in gender equality and attitudes within the economics profession.
- Our initial focus is on gender differences in doctoral fields of study in economics.

Research Objectives

1. Document and measure the extent to which there are gender differences in doctoral fields of specialization among economists.
2. Examine the roles of economic, personal, and graduate school environmental factors in determining doctoral field specialization choices.
3. Shed light on the implications of gender differences in doctoral field specialization for subsequent gender differences in salary growth and career advancement.
4. Develop a framework for examining gender differences in the evolution of fields of specialization over economists' careers.

Research Challenges

- Quantify fields of doctoral specialization in economics
- Select the degree of aggregation for field specializations
- Multi-field specialization
- Measure gender disparity in fields of specialization
- Model field specialization choices in the presence of antecedent events and sample selection: family background, K-12 experiences, decision to attend college, undergraduate experiences conditional on college attendance, decision to attend graduate school in economics, and graduate school experiences conditional on graduate school enrollment.
- Data

Research Challenges

- Data
  1. Personal background
  2. Potential salary information
  3. Graduate school characteristics

Background literature

- importance of diversity in economics and possible reasons why it is not a very diverse field (Bayer & Rouse (2016))
- gender differences in economics cannot be explained by aptitude (2014 - 43% BA in math; 28% in Economics (Emerson, McGoldrick & Mumford (2012))
- Performance of women in economics compared to other sciences (less likely to obtain tenure and it takes longer (Ginther & Kahn (2004))
- (+) correlation between female faculty share and the gender composition of the entering PhD class; lack of role models important in preventing women to be present in Economics (Hale & Regev (2014); Carrell, Page, & West (2010))
- what can be done to encourage women to enter economics (Avilova & Goldin (2018b))

Data Sources

- EconLit (graduate student information and selected sub-fields – doctoral sample)
- Academic Analytics (special arrangement through the University of Arizona – academic/salary sample)
- State websites for salary information
- *US News and World Report* – for department and university rankings

Initital conditions: field specialziation in economics at the doctoral level

Examine graduate school determinants of one's field of doctoral specialization

1. Econometrics (C)
2. Micro (D)
3. Labor/Health (I,J)
4. Macro/Finance (E,G)
5. IO (L)
6. Environ & Agric (Q)
7. Public (H)
8. Dev/Growth/Int (O,F)
9. Econ History (B,N)
10. Other (P,A,K,M,R,Y,Z)

JEL codes are aggregated into 10 mutually exclusive groupings:

1. Econometrics (C)
2. Micro (D)
3. Labor/Health (I,J)
4. Macro/Finance (E,G)
5. IO (L)
6. Environ & Agric (Q)
7. Public (H)
8. Dev/Growth/Int (O,F)
9. Econ History (B,N)
10. Other (P,A,K,M,R,Y,Z)

JEL codes are assigned in EconLit.
The code that is assigned most often determines the field of specialization.

1. 85 % of 9350 doctoral dissertations awarded during 2009-2018 report 1 main field
2. 15 % report 2 fields –1420 out of 9350
3. not included are those reporting 3 fields or more (9.5% of the original sample)

JEL codes are assigned in EconLit.
The code that is assigned most often determines the field of specialization.

1. 85 % of 9350 doctoral dissertations awarded during 2009-2018 report 1 main field <== **Our focus**
2. 15 % report 2 fields –1420 out of 9350
3. not included are those reporting 3 fields or more (9.5% of the original sample)

For J primary field choices, the individual faces J independent binary but not mutually exclusive decisions.

$$U_{ij} = Y_{ij}\theta_j + D_{ij}Z_i\gamma_j + \varepsilon_{ij}, i = 1, ...N, j = 1, ..., J$$

where U_{ij} is the utility associated with choosing field j, Y_{ij} is a vector of choice specific attributes for field j for the *ith* individual, $D_{ij} = 1(S_i = j)$, Z_i is a vector of individual (case) specific attributes including a constant term, θ_j and γ_j are parameter vectors, and ε_{ij} is a logistic distribution error term.

The model yields J independent binomial logit equations in which individuals can exhibit multiple doctoral field specializations.

The probability of each field specialization is given by

$$Prob(S_{ij} = 1) = \frac{exp(Y_{ij}\theta_j + Z_i\gamma_j)}{1+exp(Y_{ij}\theta_j + Z_i\gamma_j)}, j = 1, ..., J$$

where $S_{ij} = 1$ denotes selection of *jth* primary field, $j = 1, ..., J$ for the *ith* individual.

The individuals can choose either a single or two primary doctoral fields, thus, there are
$\frac{(J)(J-1)}{2} + J = \frac{(J)(J+1)}{2}$ possible field choice outcomes/specializations

where $\binom{J}{2} = \frac{J!}{(J-2)!\,2!} = \frac{(J)(J-1)}{2}$ is the maximum number of possible dual field specialization pairs.

and J is the number of single (primary) field specializations, .
In our case, $J = 10$ and 55 is the max number of possible dual field specialization pairs.

The independent binomial logit framework assumes away correlated field specialization choices. A more general framework is the *multivariate logit model for correlated binary choices*. Similarly to Bel et al (2018) and Russell and Petersen (2000), we implement a multivariate logit choice model with correlated choices with an assumed structure for conditional probabilities that have the same form as those from the independent binomial logit models rationalized by random utility but which allow for correlated choices:

$$Prob[S_{ij} = 1 | I_{ij}] = \frac{exp(I_{ij})}{1 + exp(I_{ij})}, \quad j = 1, ...J$$

where $I_{ij} = Y_{ij}\theta_j + Z_i\gamma_j + \sum_{l \neq j} \psi_{jl} s_{il}$, s_{ij} is the realization of the random variable S_{ij}, and $\psi_{jl} = \psi_{lj}$ is the association parameter reflecting the underlying correlation between the choices of j and l.

If we let $S_i = (S_{i1}, ..., S_{iJ})$ and $s_i = (s_{i1}, ..., s_{iJ})$, it follows from Bel et al (2018) that the joint distribution of S_i implied by the assumed conditional probabilities is given by

$$Prob(S_i = s_i | I_{i1}, ..., I_{iJ}) = \frac{exp\left[\sum_{j=1}^{J} s_{ij}(Y_{ij}\theta_j + Z_i\gamma_j) + \sum_{l>j} s_{ij}s_{il}\psi_{jl}\right]}{\sum_{m_i \in M} exp(\mu_{m_i})},$$

where M represents the choice outcome space (i.e. potentially $(J)(J+1)/2$ field choice outcome combinations).

In the estimation strategy developed in Bel et. al (2018) termed the Composite Conditional Likelihood (CCL) estimation method, the complexity of the estimation task is significantly reduced by maximum likelihood estimation of the conditional probabilities. The CCL log likelihood function is given by

$$lnL(\theta, \gamma, \psi; s) = \sum_{i=1}^{N} \sum_{j=1}^{J} \left\{ s_{ij} ln \left[\frac{exp(I_{ij})}{1 + exp(I_{ij})} \right] + (1 - s_{ij}) ln \left[\frac{1}{1 + exp(I_{ij})} \right.$$

While not fully efficient, CCL yields consistent estimators.

Conceptual Framework
Random Utility: Individual/Case Specific Variables

- Gender
- Ranking of PhD Economics Department
- Ranking of PhD Institution
- Year in which PhD Econ Degree was awarded

Field choice specific variables reflect the attributes of the choice and correspond to the outcomes for each individual under each field specialization.

Predicted (out-of-sample) relative salaries for $t = 1$ and 7 years post PhD for doctoral student i in field j, for gender $k =$ m,f

$$V_{ijk}^{(t)} = \frac{\widetilde{W}_{ijk}^{(t)}}{\sum_{l \neq j} \widetilde{W}_{ilk}^{(t)} / J},$$

where $\widetilde{W}_{ijk}^{(t)} = \exp\left(\widehat{lnW}_{ijk}^{(t)} + 0.5\hat{\sigma}_{u_{jk}}^2\right)$, and $\widehat{lnW}_{ijk}^{(t)} = X_{ijk}^{(t)}\hat{\psi}_{jk}^{(t)}$ estimated from public university salary data matched with Academic Analytics faculty data. Covariates: rankings of PhD department and PhD institution, indicator for dual fields, indicators for salary types, and years since PhD.

Conceptual Framework
Random Utility: Field Specific Variables

Predicted probabilities of academic employment for each doctoral field specialization are obtained from estimated logit models:

$$\hat{\pi}_{ijk} = \frac{exp(H_{ijk}\hat{\phi}_{jk})}{1 + (H_{ijk}\hat{\phi}_{jk})}, i = 1, ..., N_k, j = 0, ..., J, k = m, f,$$

where ϕ is a parameter vector, and H is a vector of determinants that includes indicators for the ranking of one's PhD institution and economics department, indicators for unranked institutions and economics departments, and year of PhD.

Academic employment is assigned if individual i who received their PhD in economics in year t_i appears in the PhD granting institutions sample in either year $t_i + 1$ or $t_i + 2$. The Academic Analytics database accounts for approximately 385 out 431 PhD granting universities (89%).

Gender faculty in the jth field as a share of the total faculty in the PhD granting department.

Duncan Dissimilarity Index - Primary Field Model

- The DDI for the primary field model is calculated as

$$D_{mf} = \frac{1}{2}\sum_{j=0}^{J}\left|\overline{P}_{mj} - \overline{P}_{fj}\right|, \quad 0 \leq D_{mf} \leq 1,$$

where $\overline{P}_{mj}$ and $\overline{P}_{fj}$ are the sample proportions of men and women in the jth doctoral field specialization. The value of D_{mf} represents the proportion of either gender who would have to change doctoral fields in order for there to be complete parity. $D_{mf} = 0$ implies identical field distributions for men and women, and $D_{mf} = 1$ implies complete gender segregation across fields.

- Counterfactual Duncan Indexes can be constructed using the predicted field specializations for one gender based on the parameters of the choice model estimated for the other gender.

- If the estimated choice model for men is the baseline, the Duncan index attributable to gender differences in covariates is given by

$$D_{mf}^{mc} = \frac{1}{2}\sum_{j=0}^{J}\left|\overline{P}_{mj} - \overline{P}_{fj}^{m}\right|,$$

where $\overline{P}_{fj}^{m} = \frac{\sum_{i=1}^{N_f}\hat{P}_{fij}^{m}}{N_f}$, and $\hat{P}_{fij}^{m}$ is the estimated probability that the ith female would specialize in doctoral field j if she faced the estimated male multinomial/conditional logit model.

- The value of the index arising from gender differences in the estimated choice parameters is given by

$$D_{mf}^{mp} = \frac{1}{2}\sum_{j=0}^{J}\left|\overline{P}_{fj} - \overline{P}_{fj}^{m}\right|.$$

Sample Sizes

- Salary Regressions
  1. Total: 7,193
  2. Women: 1,446
  3. Men: 5,747
- Multinomial/conditional Logit, Binomial Logit, and Academic Employment Probabilities Estimation
  1. Total: 8,853 (7,471 - 1 field)
  2. Women: 2,585 (2,204 - 1 field)
  3. Men: 6,268 (5,290 - 1 field)
- Gender Log Salary Differentials (clustered SE's by department)
  - Raw differential: -0.157 (t = -5.35)
  - Covariates (including fields of specialization): -0.080 (t = -3.06)

Empirical Results

Table 1: Difference in specialization among women and men according to 20 JEL cateogries (doctoral sample, 2009-2018).

	Women	Men	Diff.	Std. Error	P-value
General Economics and Teaching (A)	0.16	0.13	0.03	0.086	0.752
History of Ec Thought, Method. & Heter. App. (B)	0.23	0.45	-0.22	0.145	0.138
Mathematical and Quant Methods (C)	4.66	7.73	-3.07***	0.590	0.000
Microeconomics (D)	14.91	17.91	-3.01***	0.880	0.001
Macro and Monetary Econ (E)	7.69	12.37	-4.69***	0.731	0.000
International Economics (F)	7.18	6.64	0.54	0.590	0.361
Financial Economics (G)	6.68	7.99	-1.31**	0.621	0.035
Public Economics (H)	3.45	3.43	0.03	0.426	0.945
Health, Education and Welfare (I)	13.43	8.36	5.08***	0.695	0.000
Labor and Demog Economics (J)	15.92	9.99	5.93***	0.751	0.000
Law and Economics (K)	0.35	0.72	-0.37**	0.183	0.042
Industrial Organization (L)	5.82	5.94	-0.12	0.552	0.834
Business Adm, Marketing, Personnel Econ (M)	0.31	0.34	-0.03	0.134	0.849
Economic History (N)	0.70	0.80	-0.10	0.205	0.620
Ec Develop, Innovation, Tech. Change & Growth (O)	6.68	5.01	1.67***	0.533	0.002
Economic Systems (P)	0.58	0.62	-0.04	0.183	0.818
Agric & Nat Resource Economics, Envir. Ec. (Q)	7.88	8.13	-0.25	0.637	0.693
Urban, Reg, Real Estate & Transp Econ (R)	2.64	2.39	0.25	0.363	0.483
Misc. Categories (Y)	0.66	0.83	-0.17	0.206	0.403
Other: Cultural, Sociol & Anthrop Econ. (Z)	0.08	0.22	-0.15	0.100	0.142
Observations	2576	6247		DDI=13.5%	

Note: DDI – Duncan Dissimilarity Index

Difference in specialization among women and men

	Women	Men	Diff.
Econometrics (C)	3.41	5.84	-2.43***
Micro (D)	10.70	13.92	-3.22***
Labor/Health (I,J)	33.64	20.82	12.82***
Macro/Finance (E,G)	14.29	20.74	-6.45***
IO (L)	5.37	5.82	-0.45
Environ & Agric (Q)	9.06	9.52	-0.46
Public(H)	2.46	2.71	-0.25
Dev/Growth/Int (O,F)	14.16	12.86	1.30
Econ History (B,N)	0.96	1.14	-0.18
Other (P,A,K,M,R,Y,Z)	5.96	6.64	-0.67
Observations	7471		

Table: Marginal effects - multifield (doctoral sample)

	(1) Econ	(2) Micro	(3) Labor	(4) Macro	(5) IO	(6) Ag	(7) Public	(8) Dev	(9) Hist	(10) Oth
Female (d)	-0.042	-0.072	-0.341***	-0.042***	-0.459***	0.302***	-0.060**	-0.144***	0.304	0.051
	(0.074)	(0.101)	(0.004)	(0.009)	(0.147)	(0.041)	(0.029)	(0.031)	(0.442)	(0.036)

Note: Labor: Labor/Health; Macro: Macro/Finance; Ag: Agriculture & Environmental Economics;
Dev: Development/Growth/International
$N = 8{,}853; N_w = 2585\ N_m = 6268$
(d) for discrete change of dummy variable from 0 to 1

Empirical Results

	Econ	Micro	Labor/Health	Macro/Finance	IO	Agric	Public	Dev	History	Other
Women										
Share of W	−					−	−	+		−
Prob(empl)			+			−				
Rel Sal Yr 1		−	−	−	−			−		−
Rel Sal Yr 7		+	+	+	+			+		+
Share of M					−		−		−	
Men										
Share of W	−	−				−				
Prob(empl)		+	+	+	+			+	−	+
Rel Sal Yr 1	+		−		+				−	−
Rel Sal Yr 7	−		+	−					+	+
Share of M				+	−				−	

Empirical Results

	Econ	Micro	Labor/Health	Macro/Finance	IO	Agric	Public	Dev	History	Other
Women										
Econ Top 20			−		−		+			
Econ 20-40			−	−	−			+		
Econ NR			−	+	+			+		
Uni Top 20		+	+					-		+
Uni 20-40		+		+	+	−		−		
Uni NR			+			−			-	
Year		−	+	−	−	−				
Men										
Econ Top 20		−	−		+			+		−
Econ 20-40	−	−	−	−	+		+	+	-	+
Econ NR			−	−	+		+	+	+	−
Uni Top 20	−	+	+	−	+		+	−	+	+
Uni 20-40	−	+		−	+		+	−		
Uni NR		−	+	+	+	−	−	−	-	+
Year	+			+	+			+	−	

Empirical Results

	Male	Female (male coeff)	Male (female coeff)	Female
Econometrics (C)	7.56	46.46	100.00	4.45
Micro (D)	18.24	0.00	0.00	13.81
Labor/ Health (I,J)	21.12	0.00	100.00	33.38
Macro/Finance (E, G)	22.64	13.04	41.45	16.60
IO (L)	7.50	100.00	0.00	6.81
Environ & Agric (Q)	9.67	38.22	99.82	10.21
Public (H)	4.04	10.10	7.77	3.52
Dev/Growth/Int (O, F)	14.89	0.00	0.00	16.83
Econ History (B,N)	1.34	0.00	0.00	1.01
Other (P,A,K,M,R,Y,Z)	8.62	19.57	0.00	8.12
	115.60	227.39	349.04	114.74

Empirical Results

Counterfactual Duncan Dissimilarity Indexes
(DDI = 27.1)

Baseline	M_coeff/F_char characteristics	M_coeff/F_char parameters
Males	92.92	99.65
Baseline	F_coeff/M_char characteristics	F_coeff/M_char parameters
Females	98.72	95.42

Conclusions

- Women graduate students are less likely to specialize in
  - Econometrics
  - Macro
  - Agriculture and Other
- Women graduate students are more likely to specialize in
  - Labor/Health
  - Public Economics
- Trends point to somewhat increasing gender disparity in doctoral field specialization.
  - Women are under-represented in Econometrics, Micro and Macro and no evidence of positive trends for women entering these fields.
  - Women are over-represented in Labor/Health but the trends in choosing Health/Labor are positive for both women (stronger) and men.
  - negative trends in choosing Development/Growth/International for women and men and IO for women.

Conclusions

- The predictive accuracies of our model strongly dominate the naive purely random choice models.
- Field specific academic employment probabilities are generally more important than field-specific relative salaries, especially for males.
- PhD department and university rankings are factors in determining doctoral field specializations.
- Gender disparity in doctoral field specializations would be increased by only equalizing characteristics (*ceteris paribus*) or by only equalizing parameters (*ceteris paribus*).

Conclusions

- Further research will be conducted to uncover specifically which parameter and characteristics differences are driving field specialization disparity.
- A subsequent research project will seek to understand the implications of gender differences in both doctoral field specialization and post-doctoral research field specialization for gender differences among economists in academic salaries and promotions.

Thank you!
Comments/suggestions /interested in the project?
e-mail: eva.sierminska@liser.lu