The Role of Social Connections in the Labor Market

Shmuel San

The Hebrew University of Jerusalem

March 22, 2022

Introduction

- Motivating facts:
  - Some firms pay more to similar workers
  - Many/most jobs obtained through social contacts
  - Homophily of social networks
- Question: How helpful are socially connected parents for young workers' who are entering the labor market?

Literature and contributions

Effects of social connections
Importance for finding jobs (Granovetter 1973; Topa 2011); Past coworkers (Cingano and Rosolia 2012; Caldwell and Harmon 2018; Eliason et al. 2019); Parental connections (Corak and Piraino 2011; Kramarz and Skans 2014; Plug et al. 2018)
Contribution: importance of indirect parental connections

Mechanisms for the effects
Search frictions (Calvo-Armengol and Jackson 2004; Fontaine 2008); Match value: productivity (Athey et al. 2000; Bandiera et al. 2009); favoritism (Beaman and Magruder 2012; Dickinson et al. 2018), uncertainty about worker's productivity (Montgomery 1991; Dustmann et al. 2016; Bolte et al. 2020)
Contribution: separately estimate the two mechanisms

Two-sided matching models
Deterministic transferable utilities (Shapley and Shubik 1971; Demange and Gale 1985); Nondeterministic utilities (Choo and Siow 2006; Galichon and Salanié 2015)
Contribution: add search frictions (more realistic + enables simulation-based estimation)

Outline

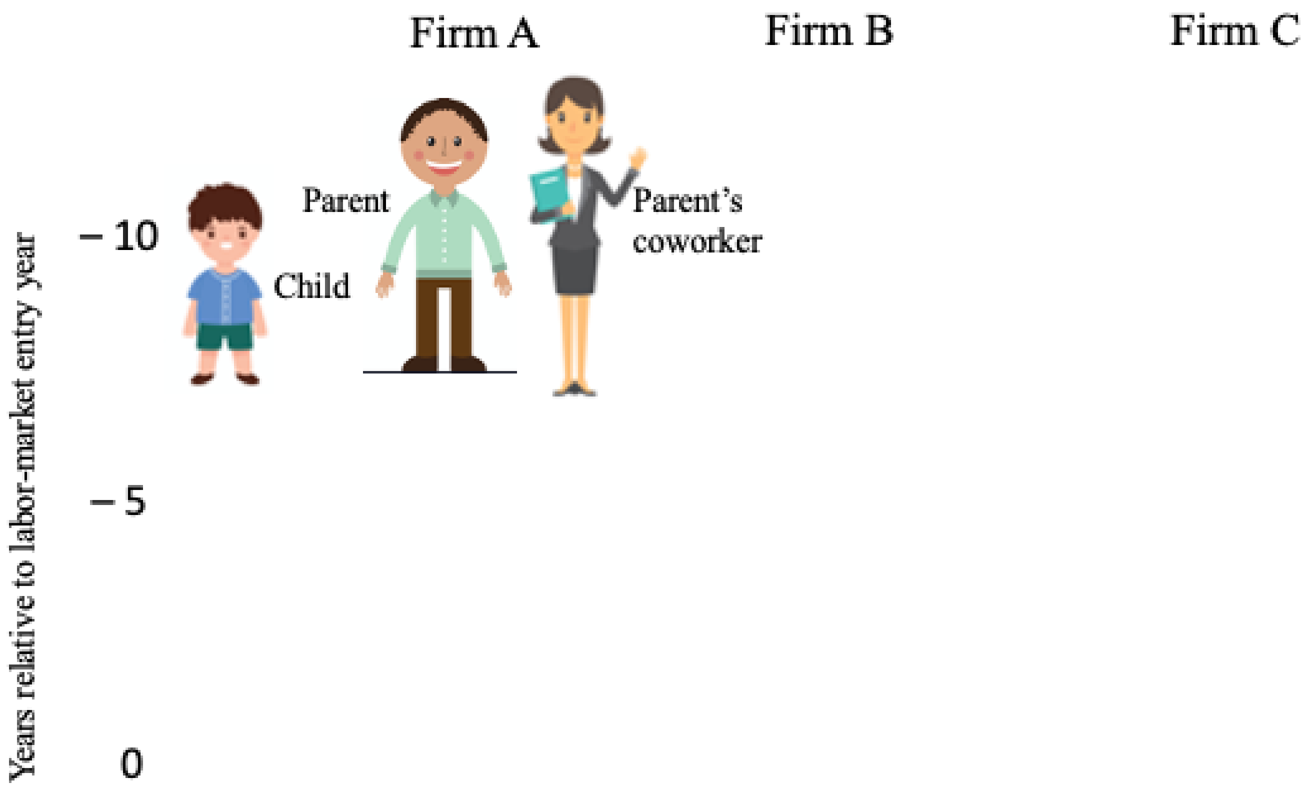
Outline

Data

- Matched employer-employee administrative records from Israel (1983-2015)
  - Person identifiers, firm identifiers, monthly indicators, yearly salary, and industry
- Israeli Population Registry
  - Date of birth, date of death, sex, ethnic group, parents identifiers, and location
- Social security records
  - Higher education (institution and years)

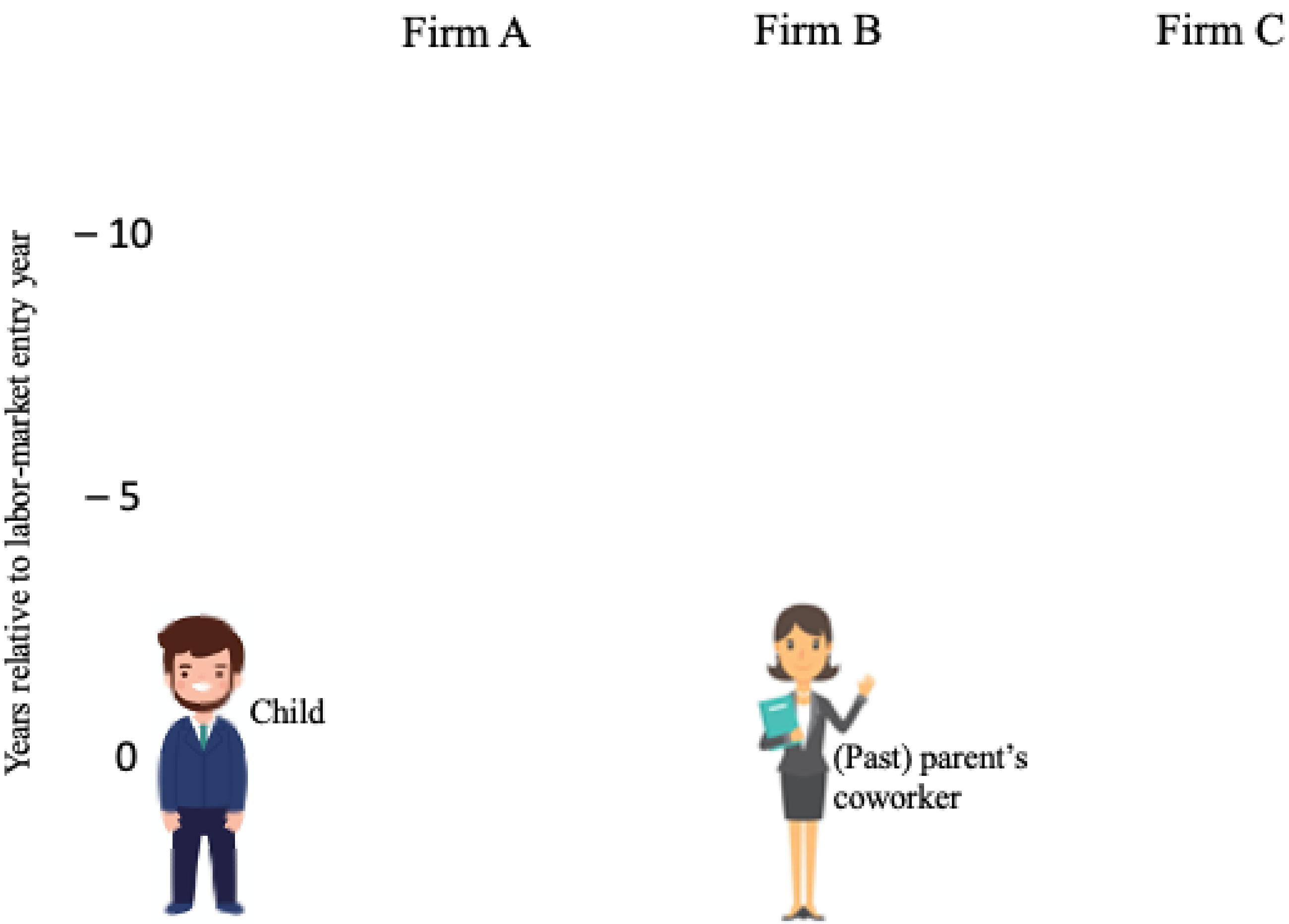
Types of parental connections

definitions



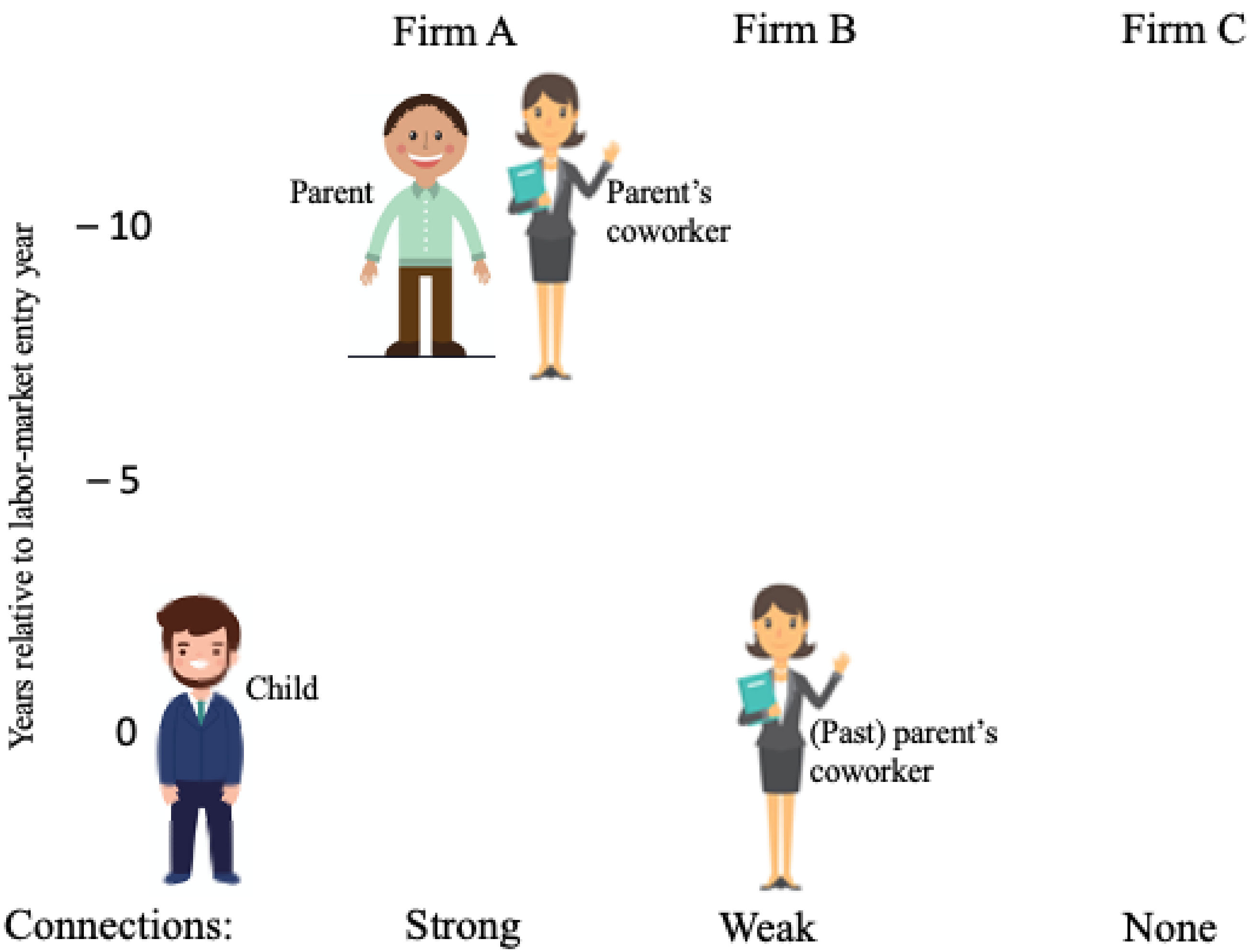
Types of parental connections

definitions



Types of parental connections

definitions



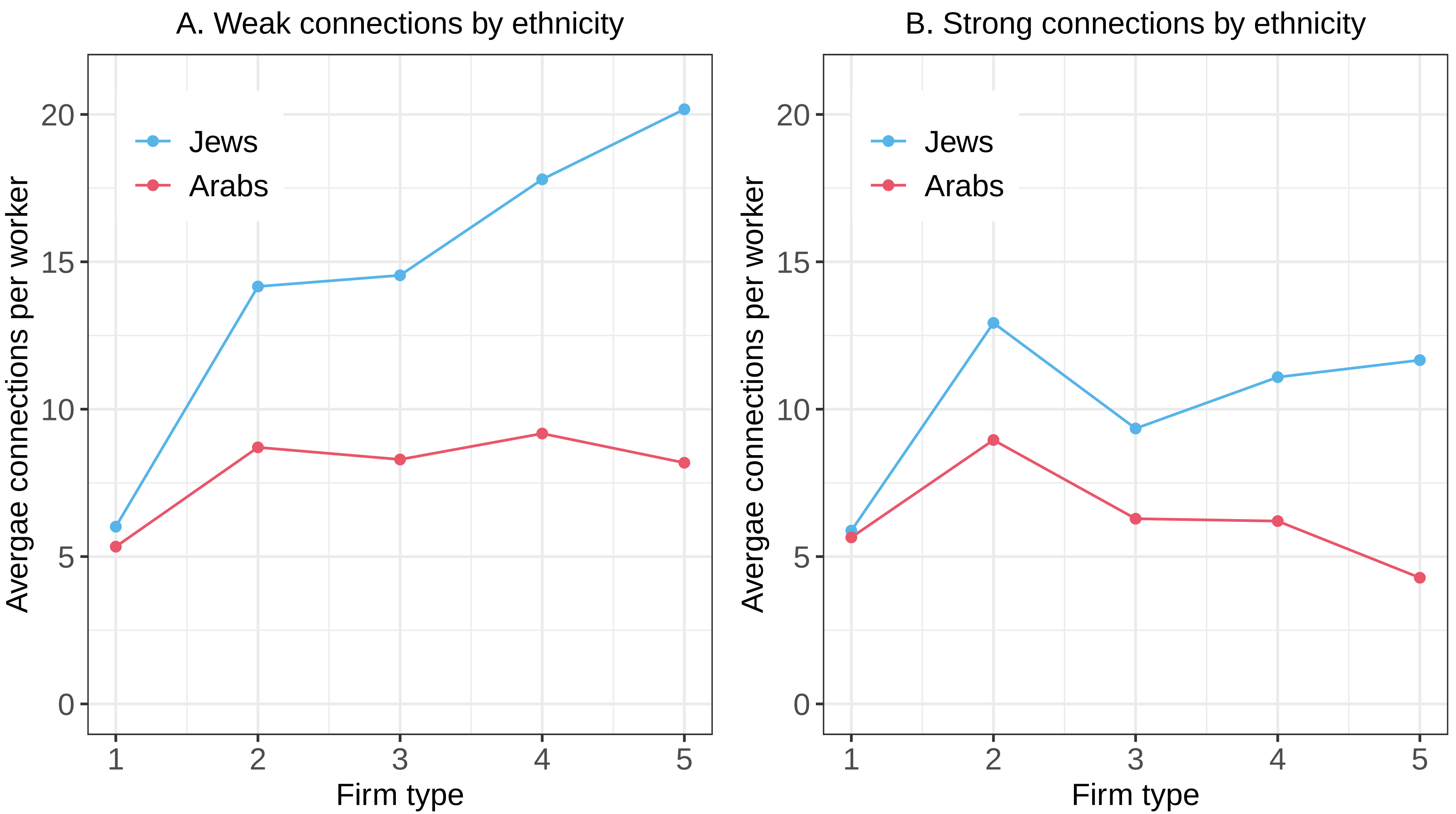
Summary statistics

Table 1: Summary statistics: new workers

	All	Ethnicity		Gender	
		Jews	Arabs	Males	Females
N.	220,806	157,023	63,783	126,233	94,573
First job					
Salary	5,839	6,053	5,312	6,223	5,325
Firm rank	0.60	0.64	0.52	0.60	0.61
Connections					
Weak	0.03	0.02	0.04	0.03	0.02
Strong	0.11	0.09	0.17	0.13	0.08
Connections quality					
Av. firm rank					
Weak	0.64	0.66	0.58	0.63	0.65
Strong	0.61	0.64	0.54	0.60	0.62

sample selection AKM wage regression

Connections per worker by ethnicity

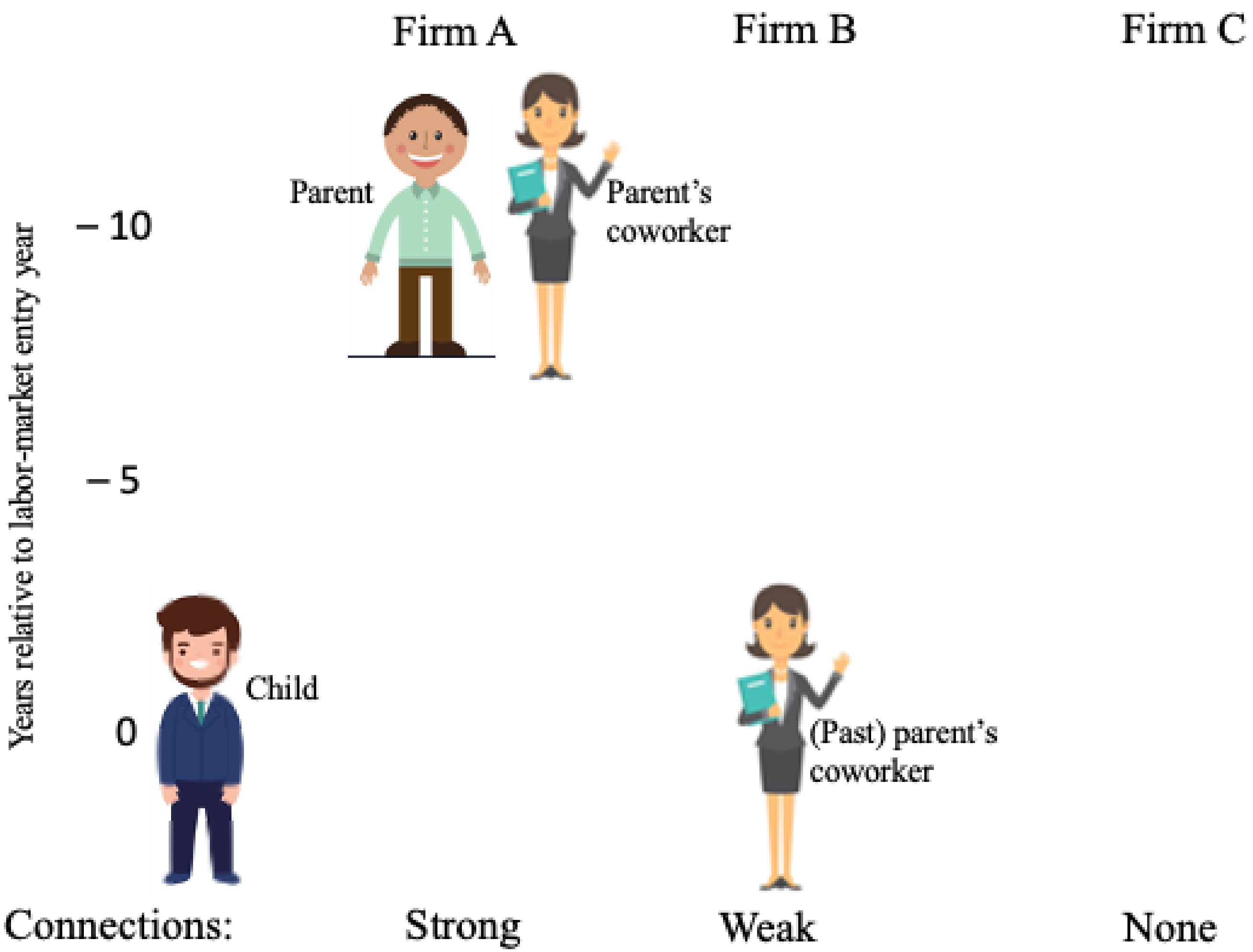


Gender

Outline

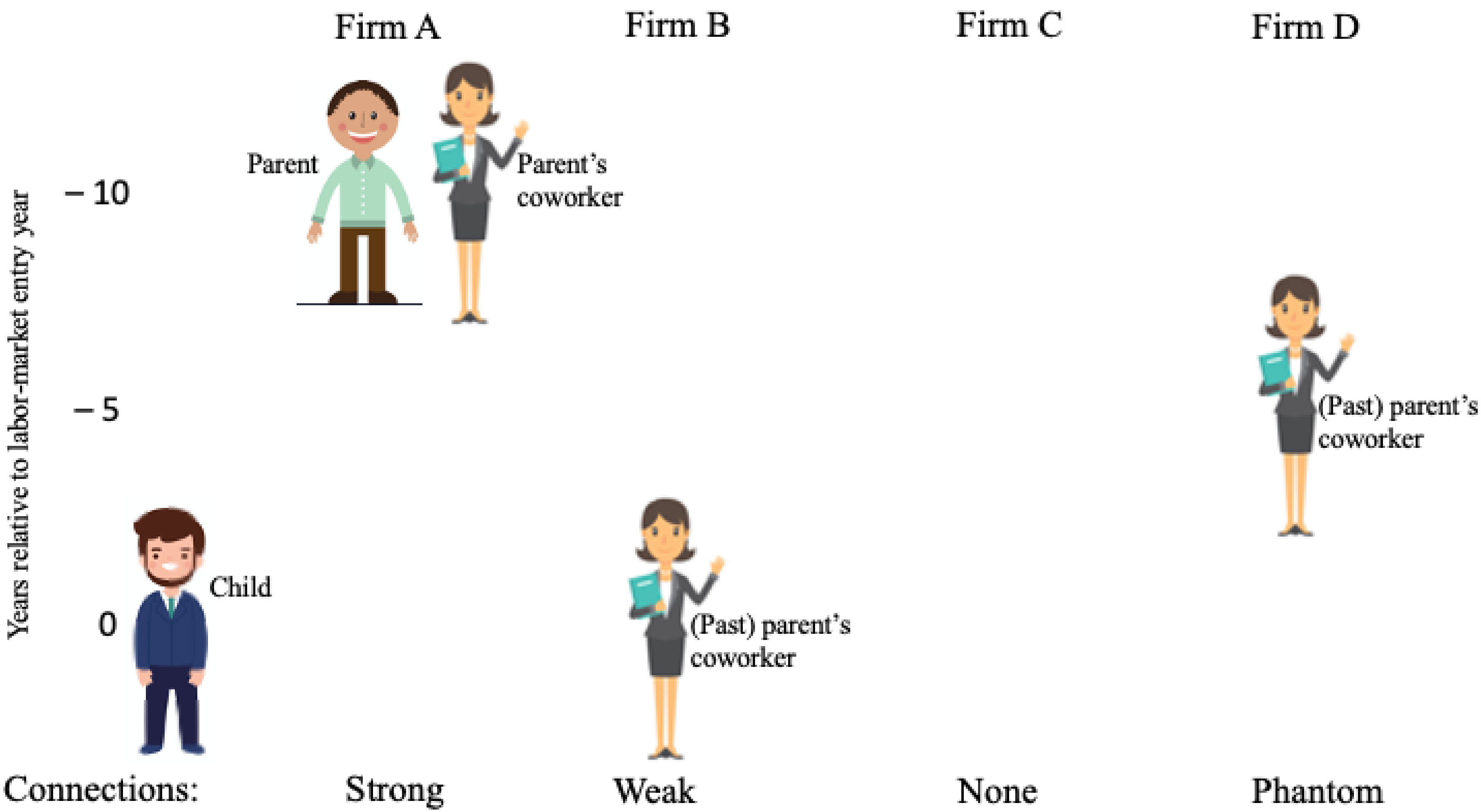
Types of parental connections

definitions

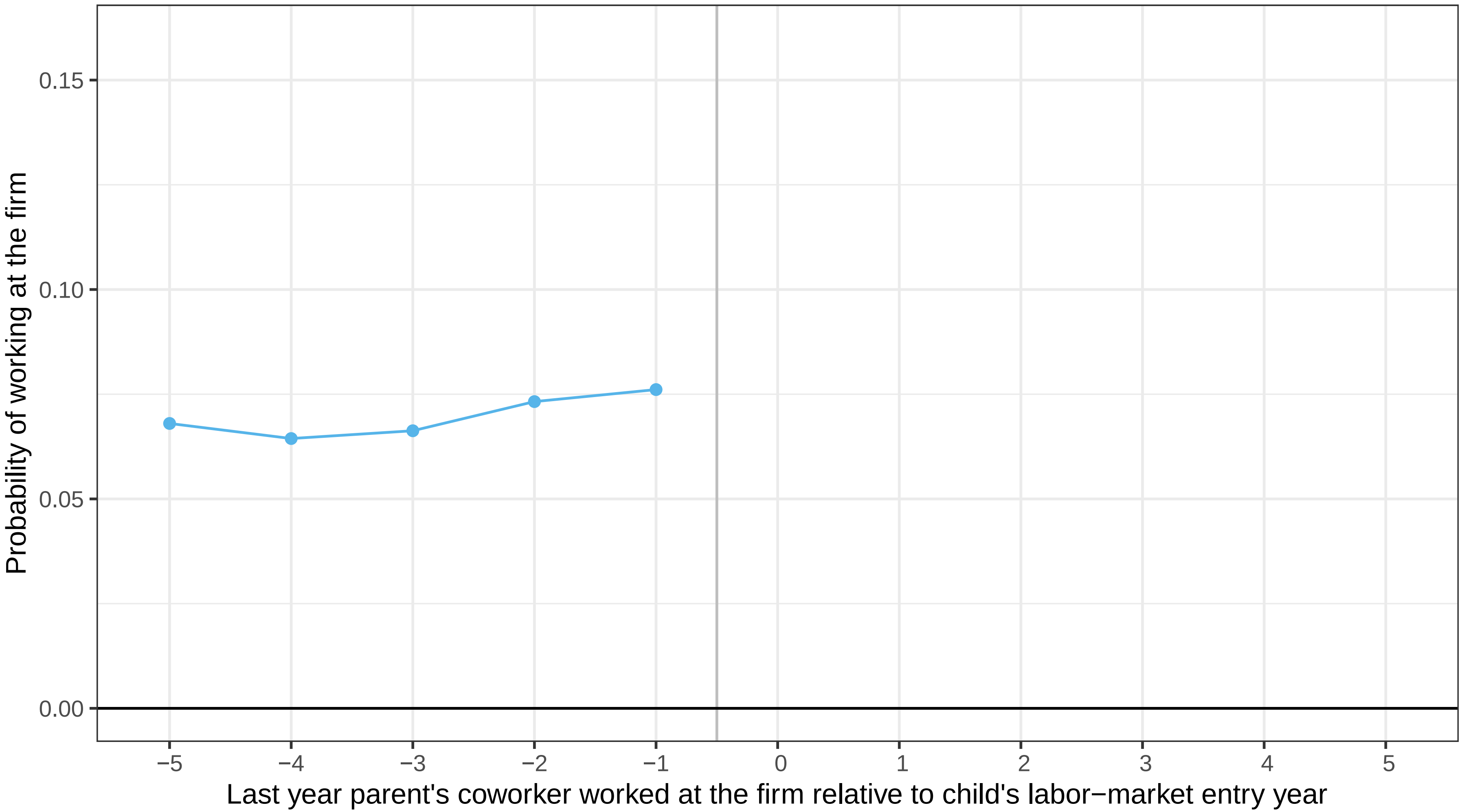


Types of parental connections

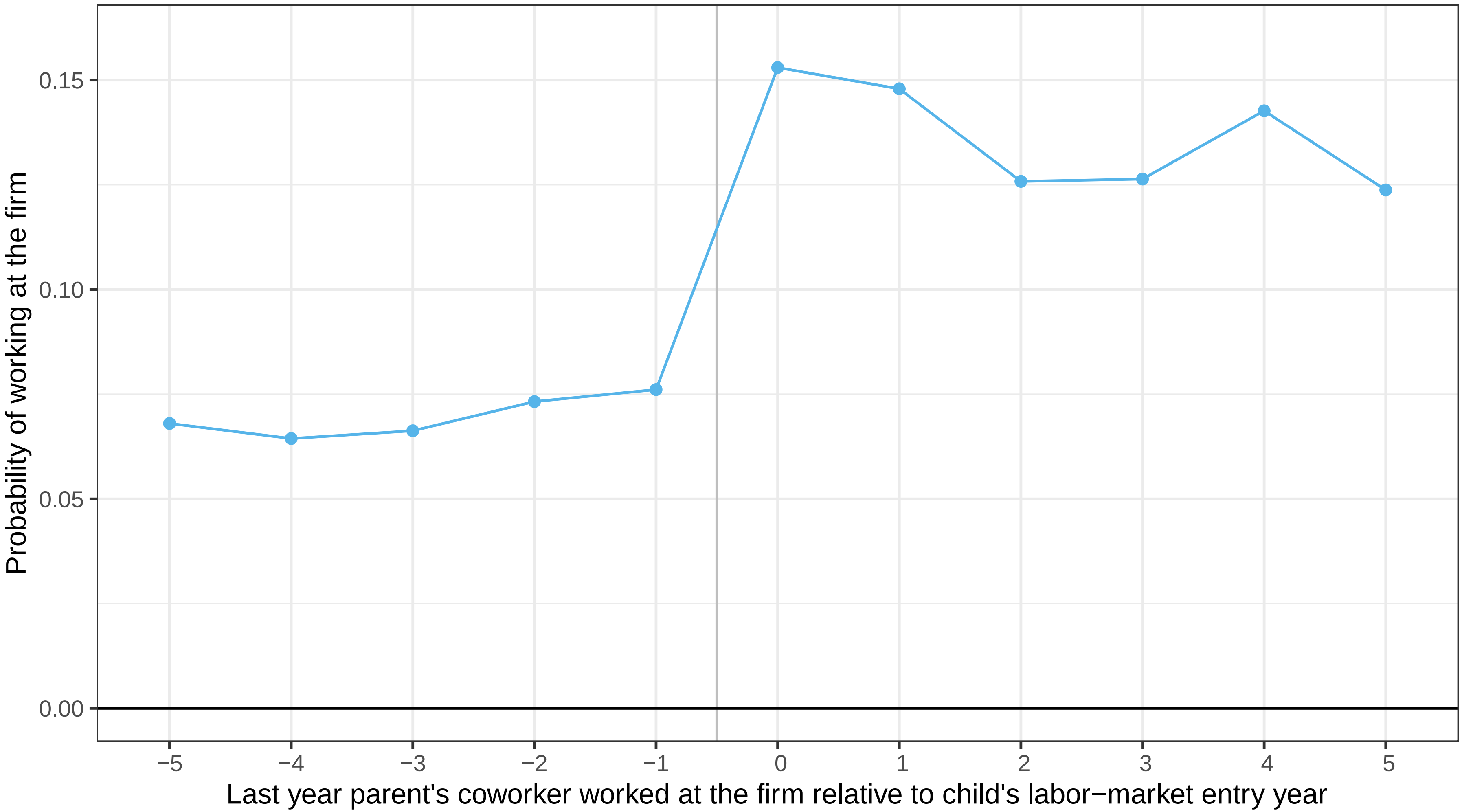
definitions balancing table



Employment probability: raw data



Employment probability: raw data



Econometric model

- Extending Kramarz and Skans (2014) fixed-effects transformation framework
- Group workers based on observables
- The probability that a worker i of a group x starts working in firm j is

$$e_{ixj} = \phi_{xj} + \sum_{c=p,w,s} \delta^c \cdot D^c_{ij} + \epsilon_{ixj}$$

with

- $e_{ixj} = 1$ if i worked at firm j
- ϕ_{xj} group-firm match specific effect
- $D^c_{ij} = 1$ if i had connections of type c at firm j

Within-group estimation in practice

- Restrict the sample to cases where there is within group-firm variation in $D_{ij} \equiv \max_c D_{ij}^c$
- For each group-firm combination, compute
  - The fraction of connected children who were hired by the firm

$$R_{xj}^{CON} = \frac{\sum_{i \in x} e_{ixj} D_{ij}}{\sum_{i \in x} D_{ij}} = \phi_{xj} + \sum_{c=1}^{C} \delta^c \cdot D_{xj}^c + \epsilon_{xj}^{CON}$$

  - The fraction of non-connected children who were hired by firm j

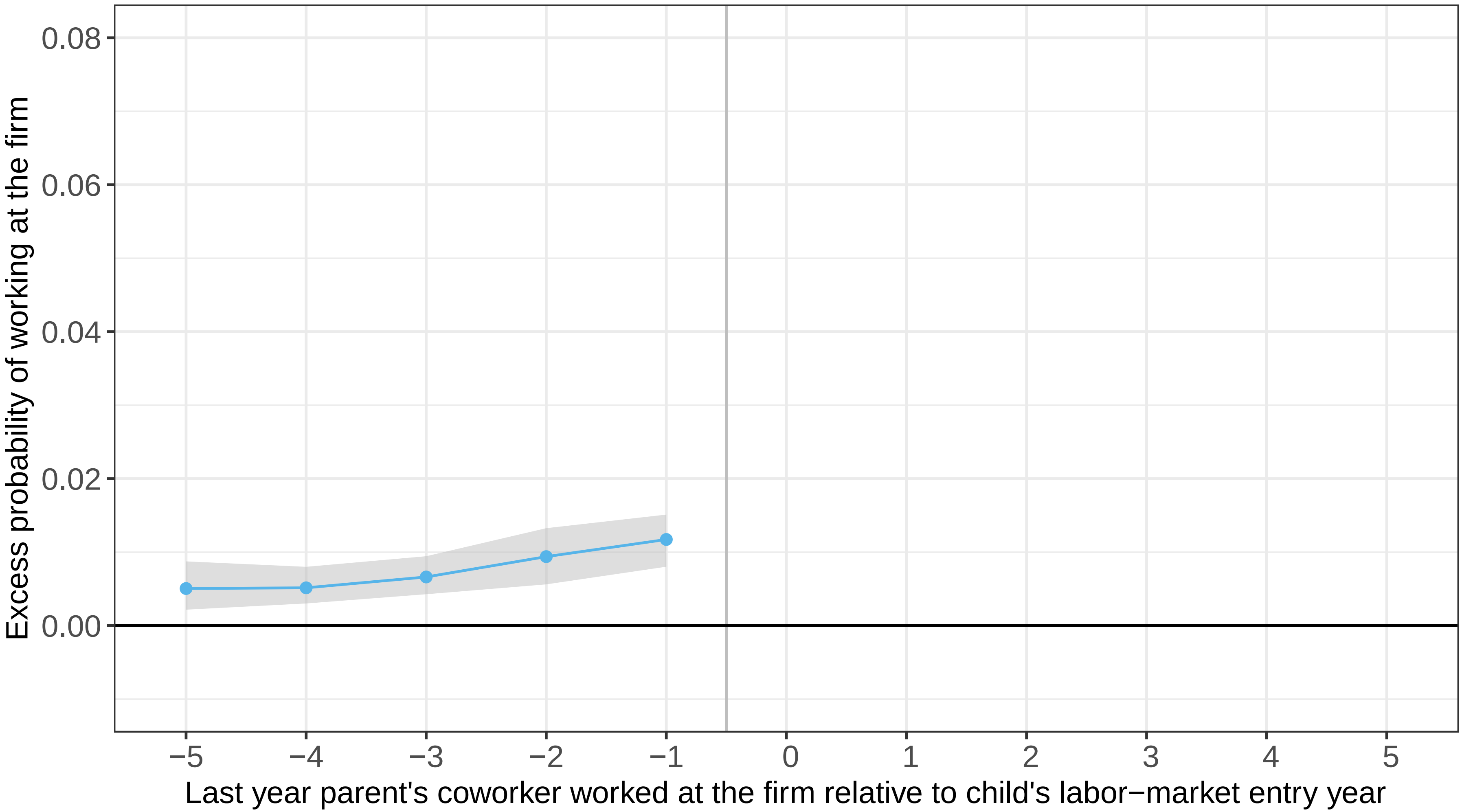
$$R_{xj}^{-CON} = \frac{\sum_{i \in x} e_{ixj}(1 - D_{ij})}{\sum_{i \in x}(1 - D_{ij})} = \phi_{xj} + \epsilon_{xj}^{-CON}$$

- Estimate

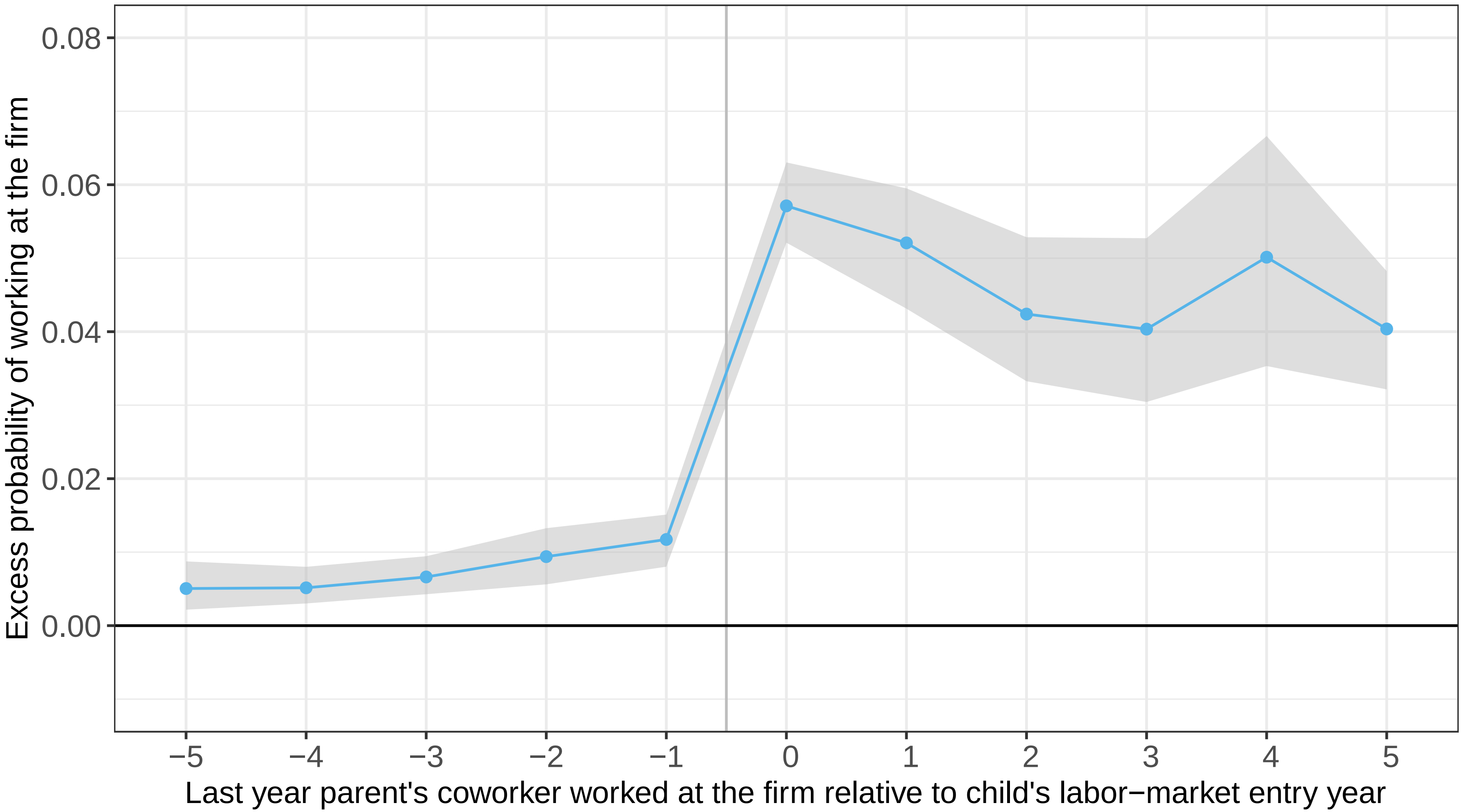
$$R_{xj} \equiv R_{xj}^{CON} - R_{xj}^{-CON} = \sum_{c=1}^{C} \delta^c \cdot D_{xj}^c + \epsilon_{xj}^{G}.$$

Outline

Effects of connections on employment: Event study



Effects of connections on employment: Event study



Effects of connections on employment: Average effects

Table 2: Effects of parental connections on firm assignment

	All	Jews	Arabs	Males	Females
	(1)	(2)	(3)	(4)	(5)
Phantom connections	0.010	0.006	0.030	0.011	0.008
	[0.009,0.011]	[0.005,0.007]	[0.025,0.032]	[0.010,0.013]	[0.006,0.010]
Weak connections	0.050	0.031	0.143	0.067	0.031
	[0.047,0.054]	[0.028,0.034]	[0.131,0.156]	[0.061,0.071]	[0.027,0.036]
Strong connections	0.487	0.366	0.917	0.617	0.338
	[0.472,0.501]	[0.351,0.384]	[0.878,0.956]	[0.593,0.647]	[0.320,0.354]
R0 (no connections)	0.005	0.005	0.006	0.005	0.006
	[0.005,0.005]	[0.005,0.005]	[0.006,0.006]	[0.005,0.005]	[0.005,0.006]
Ratio weak-phantom	3.666	3.259	4.177	4.409	2.731
	[3.316,4.081]	[2.841,3.681]	[3.651,4.803]	[3.912,4.959]	[2.262,3.303]
Ratio strong-phantom	32.52	33.99	25.91	38.37	25.37
	[30.02,35.53]	[30.65,37.8]	[23.52,30.03]	[34.83,43.67]	[22.41,29.39]
Observations	21,166,443	16,837,526	4,328,917	15,319,313	5,847,130
N firms	149,729	144,186	117,746	145,939	134,555
N groups	2,959	1,658	1,301	1,548	1,411
N workers	220,684	157,009	63,675	170,872	49,812
N connections	40,827,833	33,261,814	7,566,019	31,664,340	9,163,493

Robustness checks

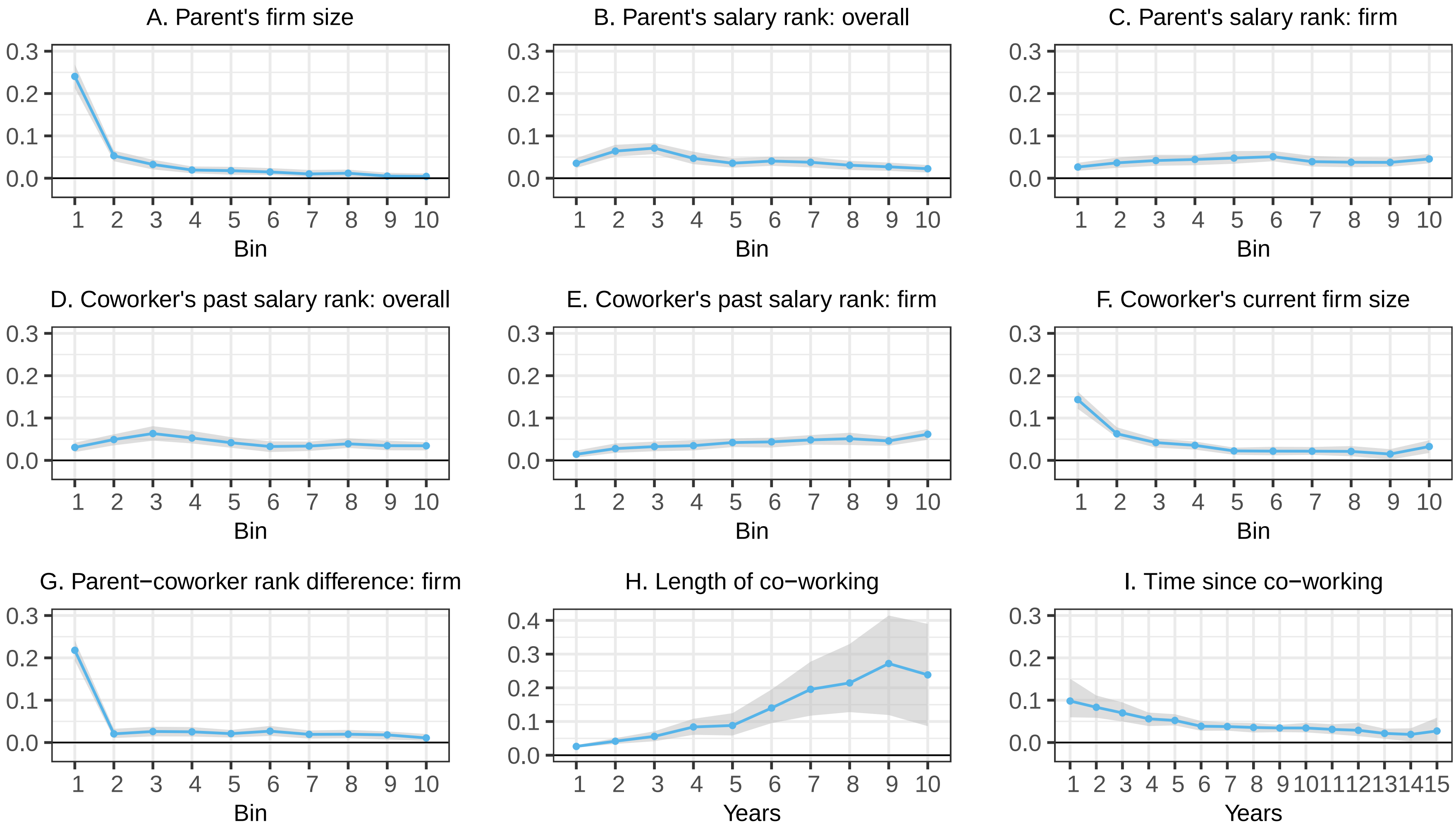
- Exogenous separations (death and retirement of contacts) go
- Placebo connections go
- Definitions of connections go

Heterogeneity of the effect

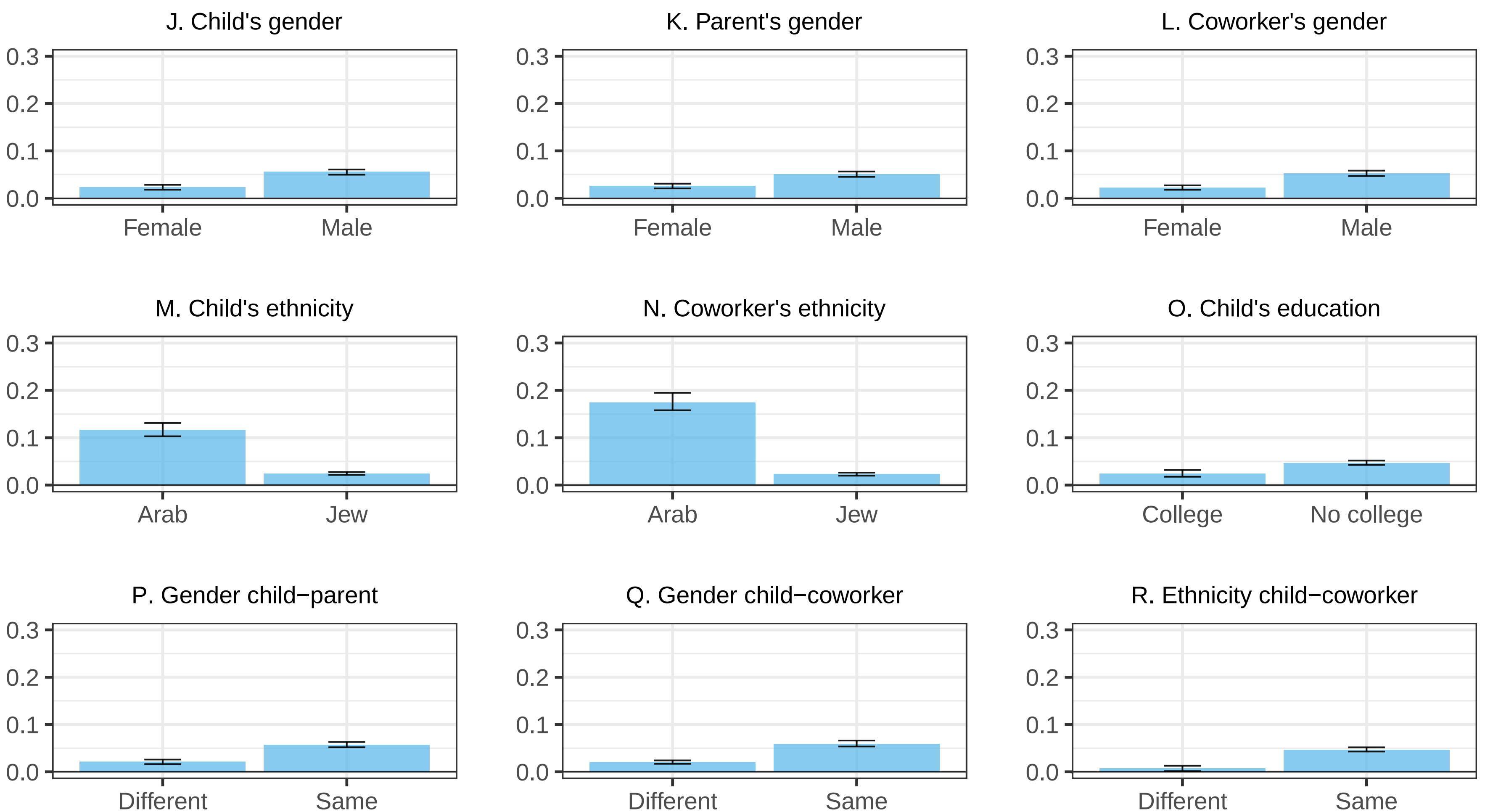
- Dividing phantom and weak connections into disjoint sets based on characteristics of the workers and the connections

$$
\begin{aligned}
e_{ixj} = & \alpha_{xj} + \sum_{c'} \left(\delta^{w,c'} \cdot D^{w,c'} + \delta^{p,c'} \cdot D^{p,c'} \right) + \\
& \delta^{s} \cdot D^{s}_{ij} + \epsilon_{ixj}
\end{aligned}
$$

Heterogeneity (1/2)



Heterogeneity (2/2)



stylized facts

Correlation with salary

- Correlation between connections at first job and salary

$$w_i = \sum_{c=p,w,s} \delta^c D^c_{i,j(i)} + \phi_{x(i)} + \psi_{j(i)} + \epsilon_i.$$

where

- $j(i)$ is the firm in which i works at
- $x(i)$ is the observable group of worker i (ethnicity, education, gender, year of first job, age, district)
- $D^c_{i,j}$ indicates connection of type c between i and j

- This analysis does not identify the causal effect: ignores selection

Salary and tenure at first job

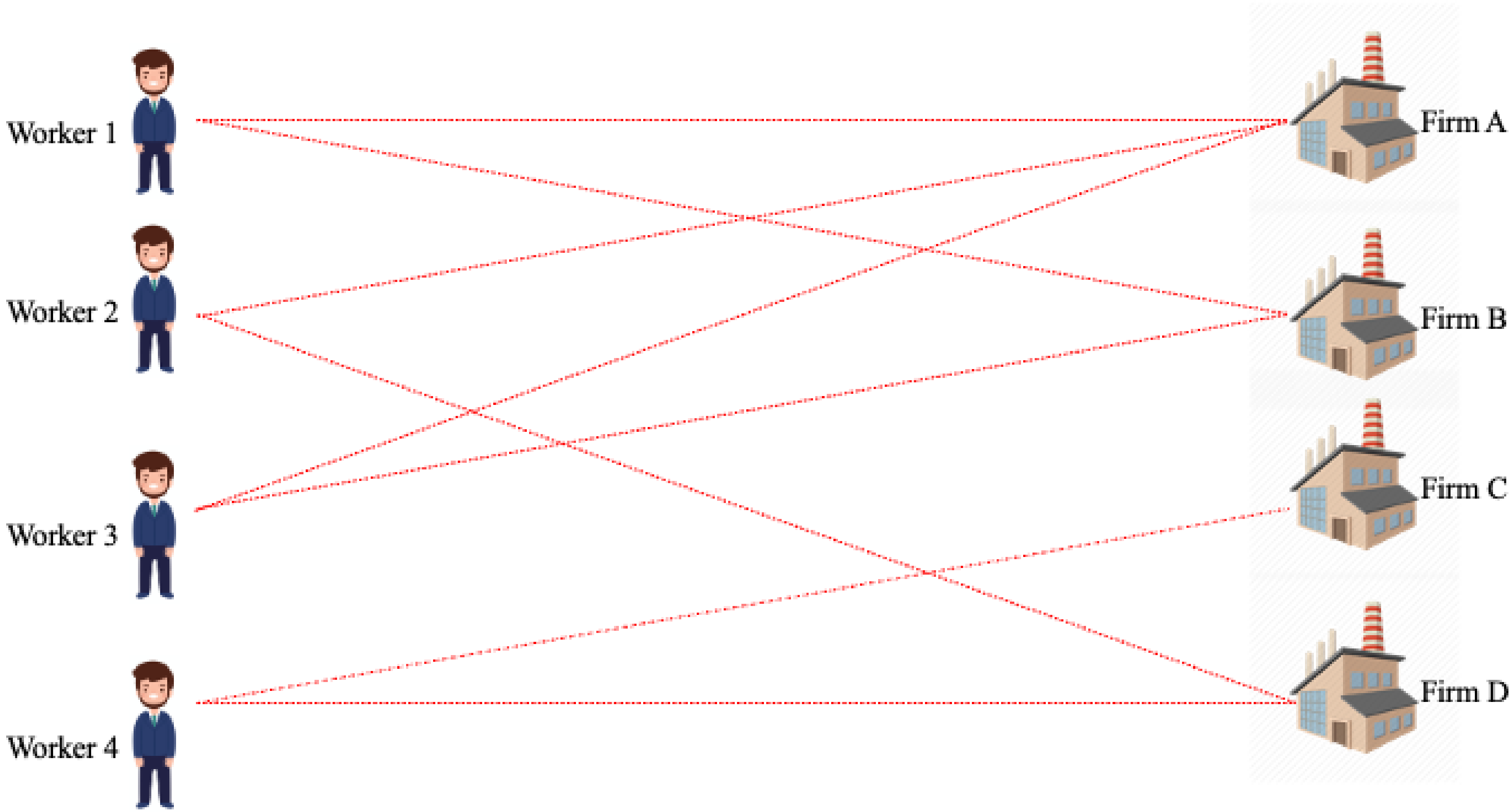
Table 3: Correlation between parental connections at first job and salary and tenure

	Log salary		Job tenure	
	(1)	(2)	(3)	(4)
Phantom connections	-0.007	0.012	0.123	0.098
	(0.005)	(0.004)	(0.022)	(0.022)
Weak connections	0.018	0.026	0.182	0.187
	(0.005)	(0.004)	(0.024)	(0.025)
Strong connections	0.074	0.083	0.601	0.441
	(0.004)	(0.003)	(0.024)	(0.020)
Group FE	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Observations	220,806	220,806	220,806	220,806
N firms	54,321	54,321	54,321	54,321
R^2 (full model)	0.169	0.624	0.127	0.414
R^2 (projected model)	0.004	0.006	0.014	0.007

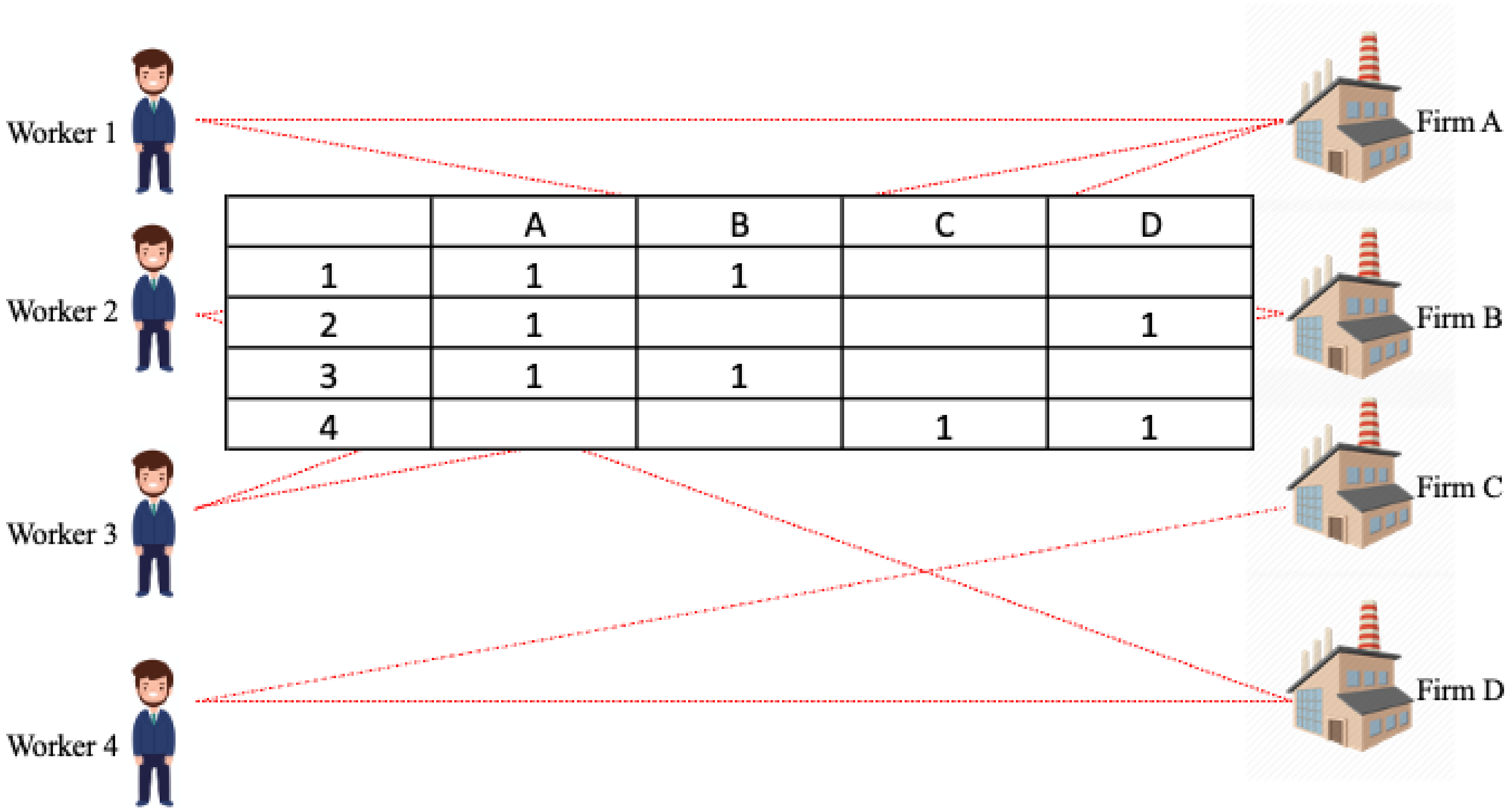
$$w_i = \sum_{c=1}^{C} \delta^c D^c_{i,j(i)} + \phi_{x(i)} + \psi_{j(i)} + \epsilon_i.$$

Outline

Connections: RHS (data)

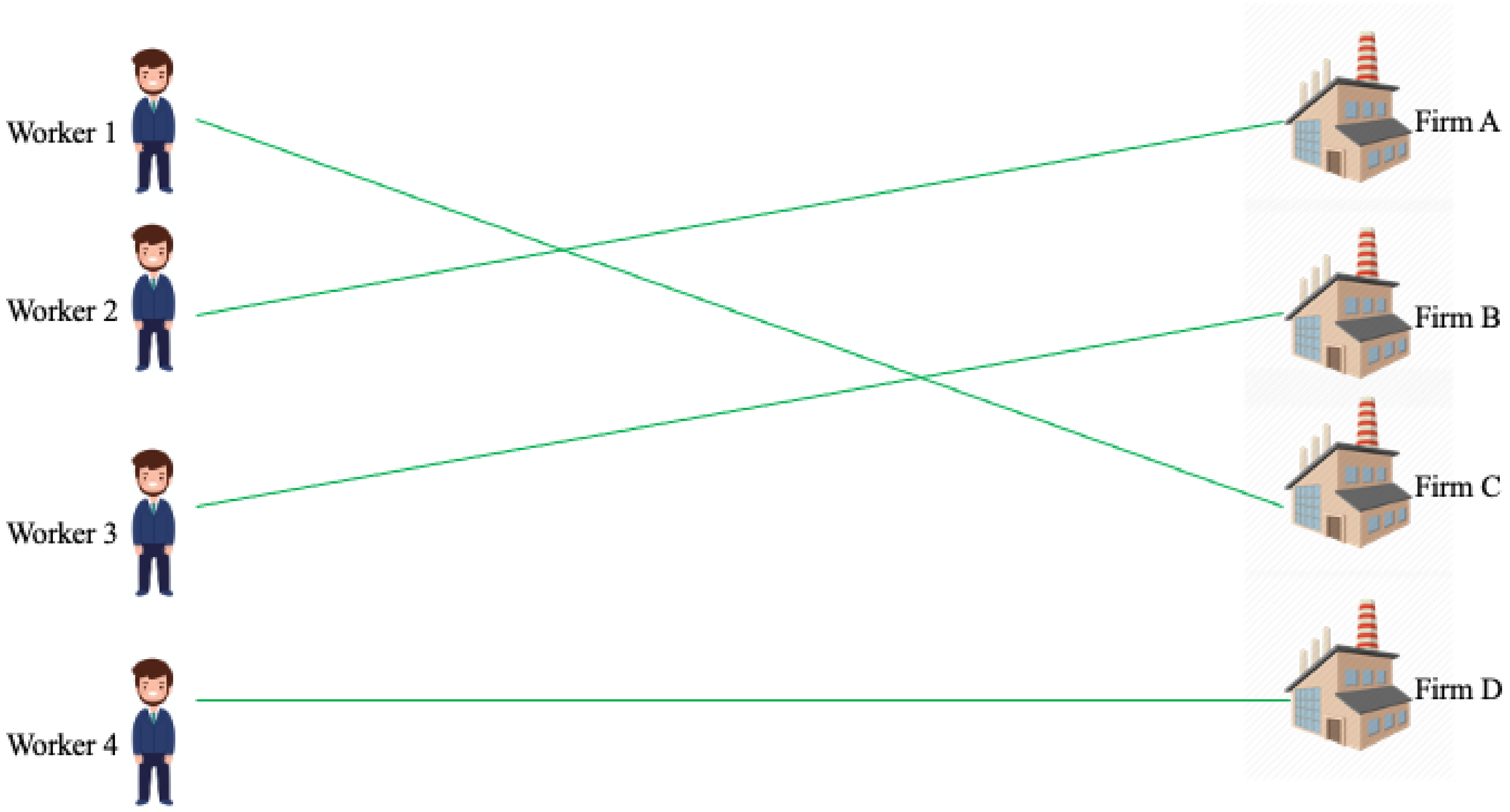


Connections: RHS (data)

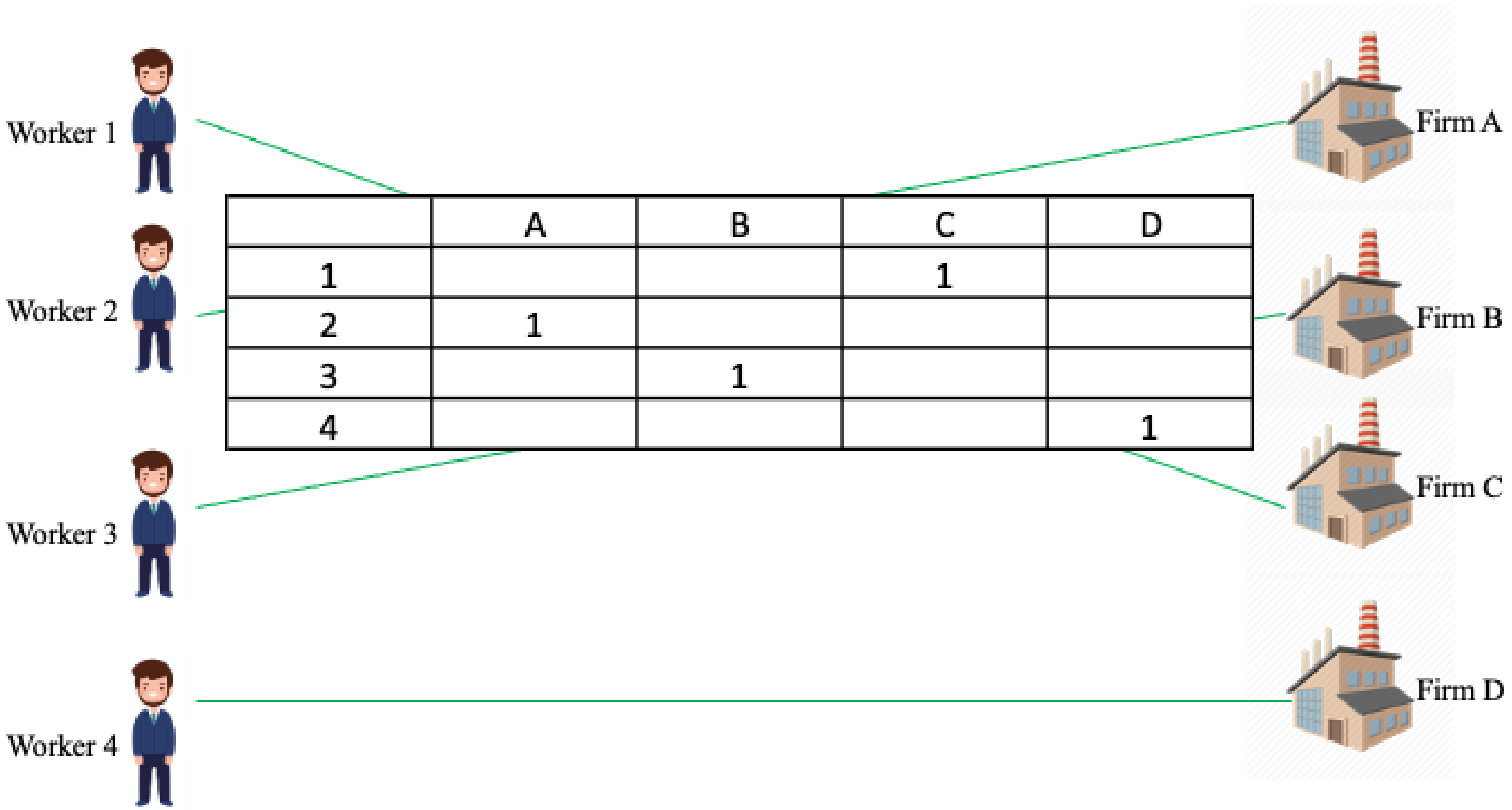


	A	B	C	D
1	1	1		
2	1			1
3	1	1		
4			1	1

Matches: LHS 1 (data)



Matches: LHS 1 (data)

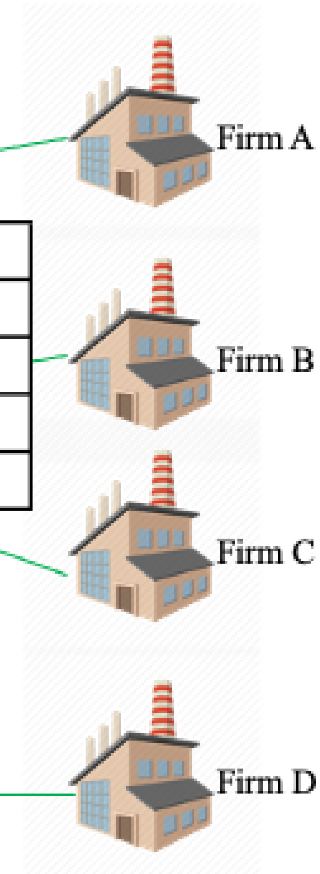


	A	B	C	D
1			1	
2	1			
3		1		
4				1

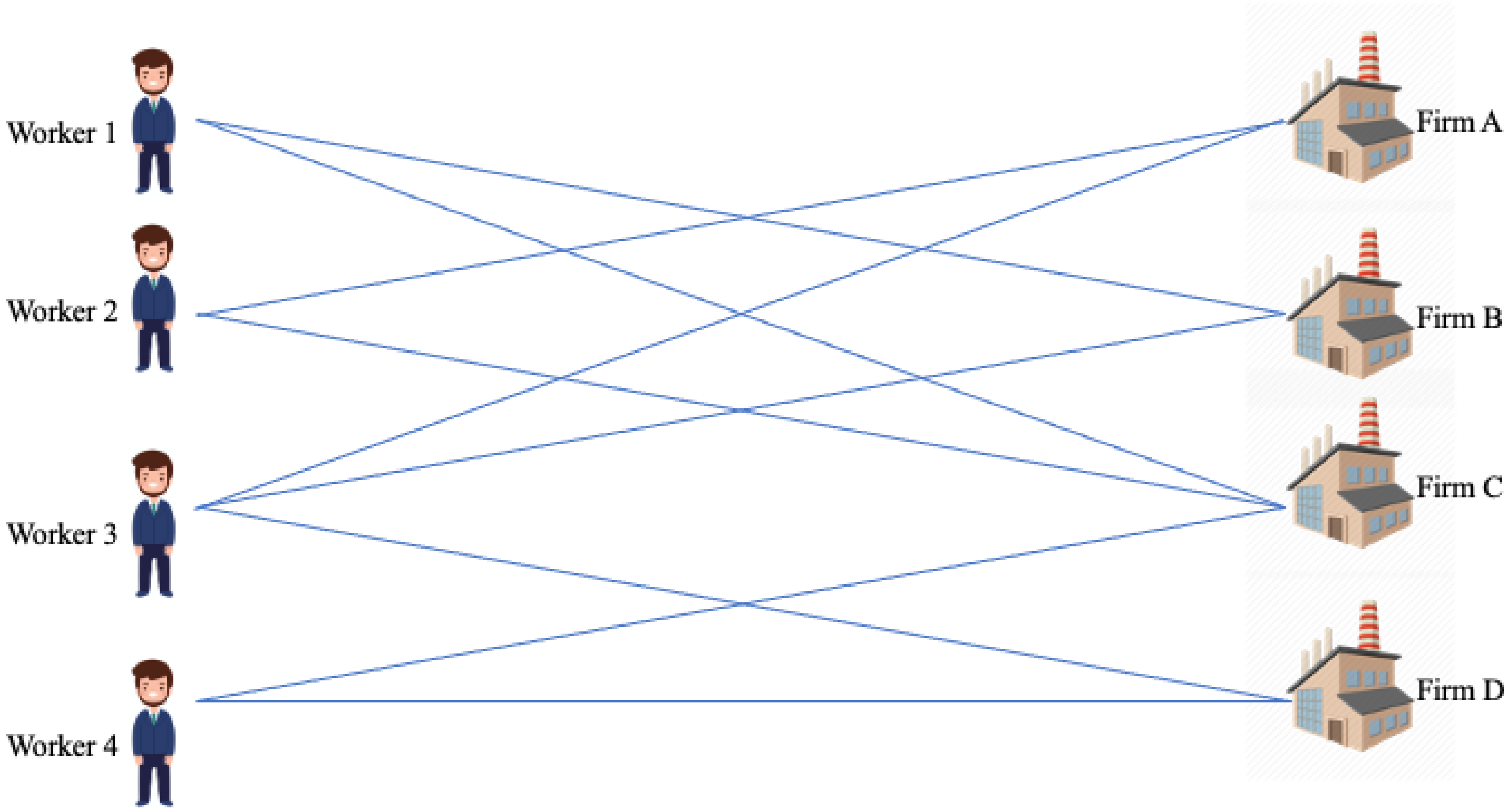
Wages: LHS 2 (data)

	A	B	C	D
1			1	
2	1			
3		1		
4				1

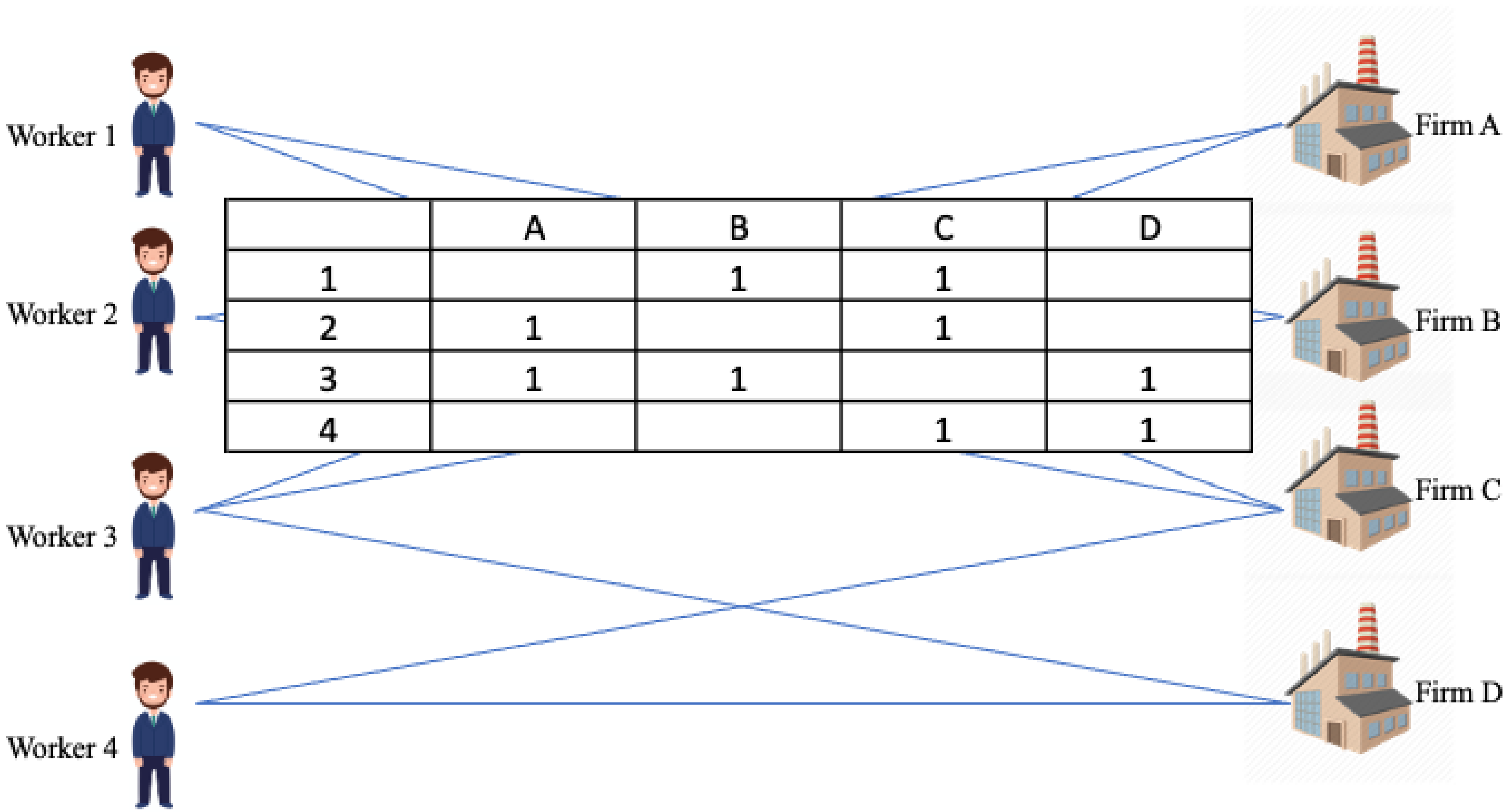
	Wages
1	5
2	3
3	2
4	7



Meetings: parameter 1 (model)



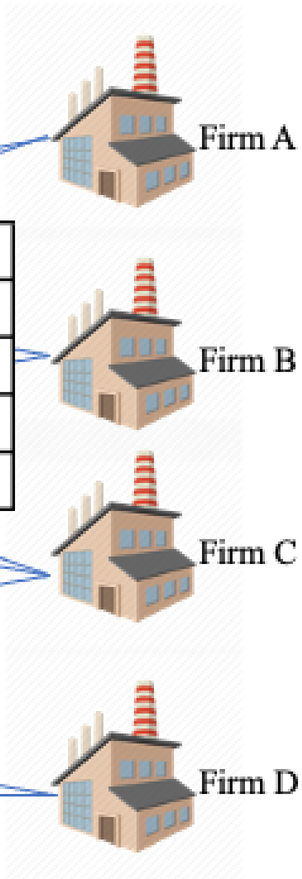
Meetings: parameter 1 (model)



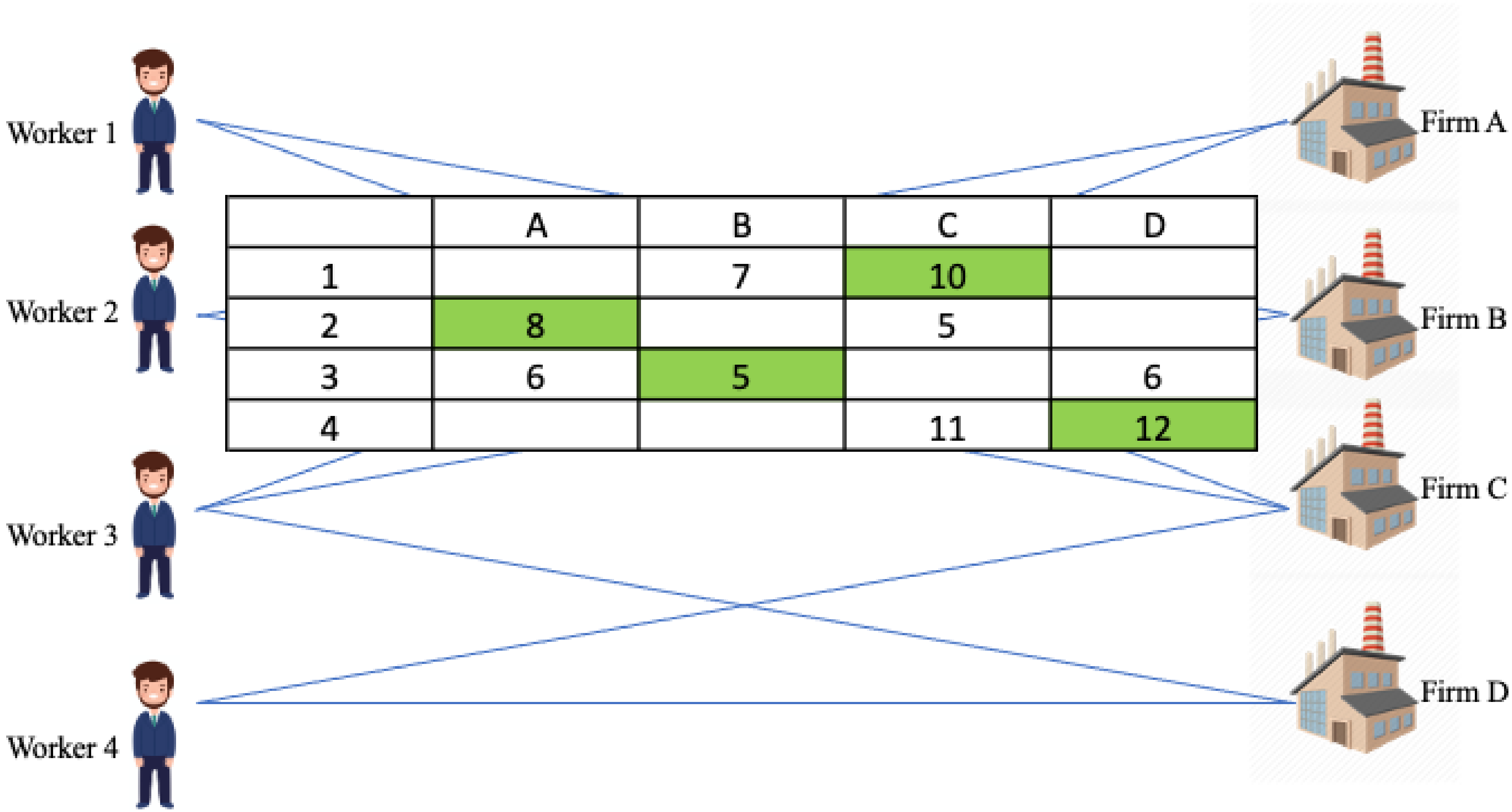
	A	B	C	D
1		1	1	
2	1		1	
3	1	1		1
4			1	1

Match utility: parameter 2 (model)

	A	B	C	D
1		7	10	
2	8		5	
3	6	5		6
4			11	12

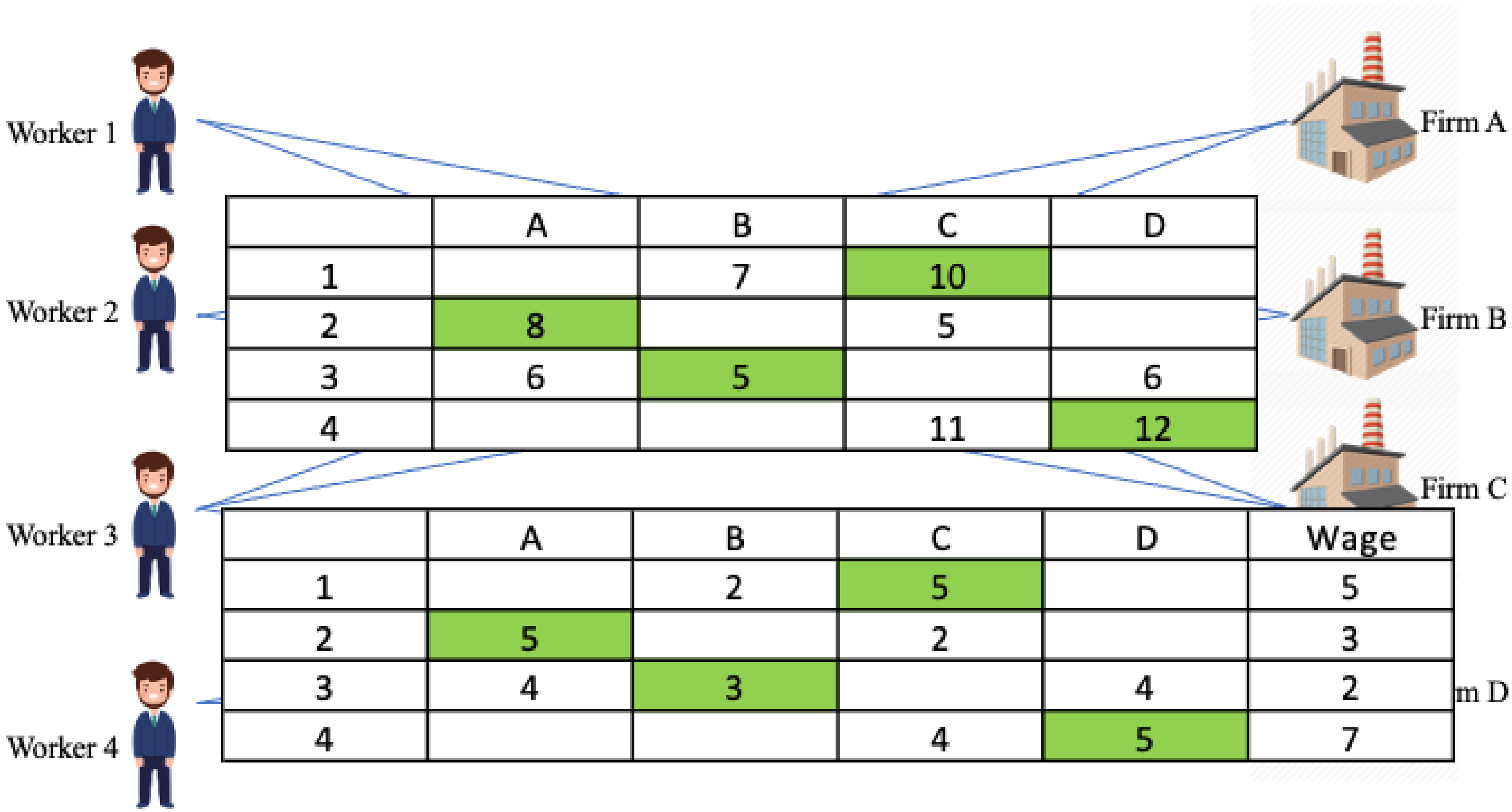


Equilibrium matches: prediction 1



	A	B	C	D
1		7	10	
2	8		5	
3	6	5		6
4			11	12

Equilibrium wages: prediction 2



	A	B	C	D
1		7	10	
2	8		5	
3	6	5		6
4			11	12

	A	B	C	D	Wage
1		2	5		5
2	5		2		3
3	4	3		4	2
4			4	5	7

Set-up

- X types of workers, Y types of firms
- T markets
- In each market t, I_t workers, J_t firms (jobs), $I_t = J_t$, I_{tx} workers of type $x \in \mathcal{X}$, J_{ty} firms of type $y \in \mathcal{Y}$
- Each worker i and firm j are connected by exactly one type of connection $c = 0, 1, ..., C$
- Matching in two stages:
    - Workers and firms randomly meet
    - Given meetings: each worker chooses the best firm and vice versa; wages clear the markets

Stage 1: meeting

- The meeting probability depends on the observable characteristics of i and j

$$m_{ij} = 1\left(\rho_{ij} \leq p_{txyc}\right)$$

  - m_{ij}: meeting indicator
  - ρ_{ij}: iid standard uniform
  - p_{txyc}: systematic meeting probability

Stage 2: matching

- After the realization of the meetings, there is a matching process between all feasible pairs
- Transferable utilities (TU)
- The utility of a firm j which employs a worker i is:

$$V_{ij} = f_{ij} - w_{ij}$$

where

$$log(f_{ij}) = b + \beta_{txyc} + \sigma \cdot \xi_{ij} \ , \ \xi_{ij} \sim N(0,1)$$

- The utility of the worker is:

$$U_{ij} = w_{ij}$$

equilibrium definition

Equilibrium characterization: matching

- Equilibrium matching is generically unique
- (Shapley and Shubik 1971): μ is an equilibrium matching if and only if it maximizes the total joint surplus $f_{ij} = U_{ij} + V_{ij}$

$$\mu \in \operatorname{argmax}_{\mu'} \sum_{\mu'(i,j)=1} f_{ij}$$
$$\text{s.t.} \quad \mu' \text{ is feasible}$$

- Equilibrium matching can be found efficiently using the auction algorithm (Bertsekas 1998) auction algorithm

Equilibrium characterization: wages

- Equilibrium wages are not unique
- If w is an equilibrium wage schedule, so is $w + r$
- The set of (normalized) equilibrium wages is a lattice: there exist $\{\underline{w}_i, \bar{w}_i\}_{i=1}^{I}$ such that $\{w_i | \underline{w}_i \leq w_i \leq \bar{w}_i\}_{i=1}^{I}$ is the set of equilibrium wages (Demange and Gale 1985)
- Find the bounds using the Bellman-Ford algorithm (Bonnet et al. 2018) BF algorithm example
- Wages are $w_i = (1 - \lambda)\underline{w}_i + \lambda \bar{w}_i$ for some "bargaining power" $\lambda \in [0, 1]$

summary (inner loop)

Outline

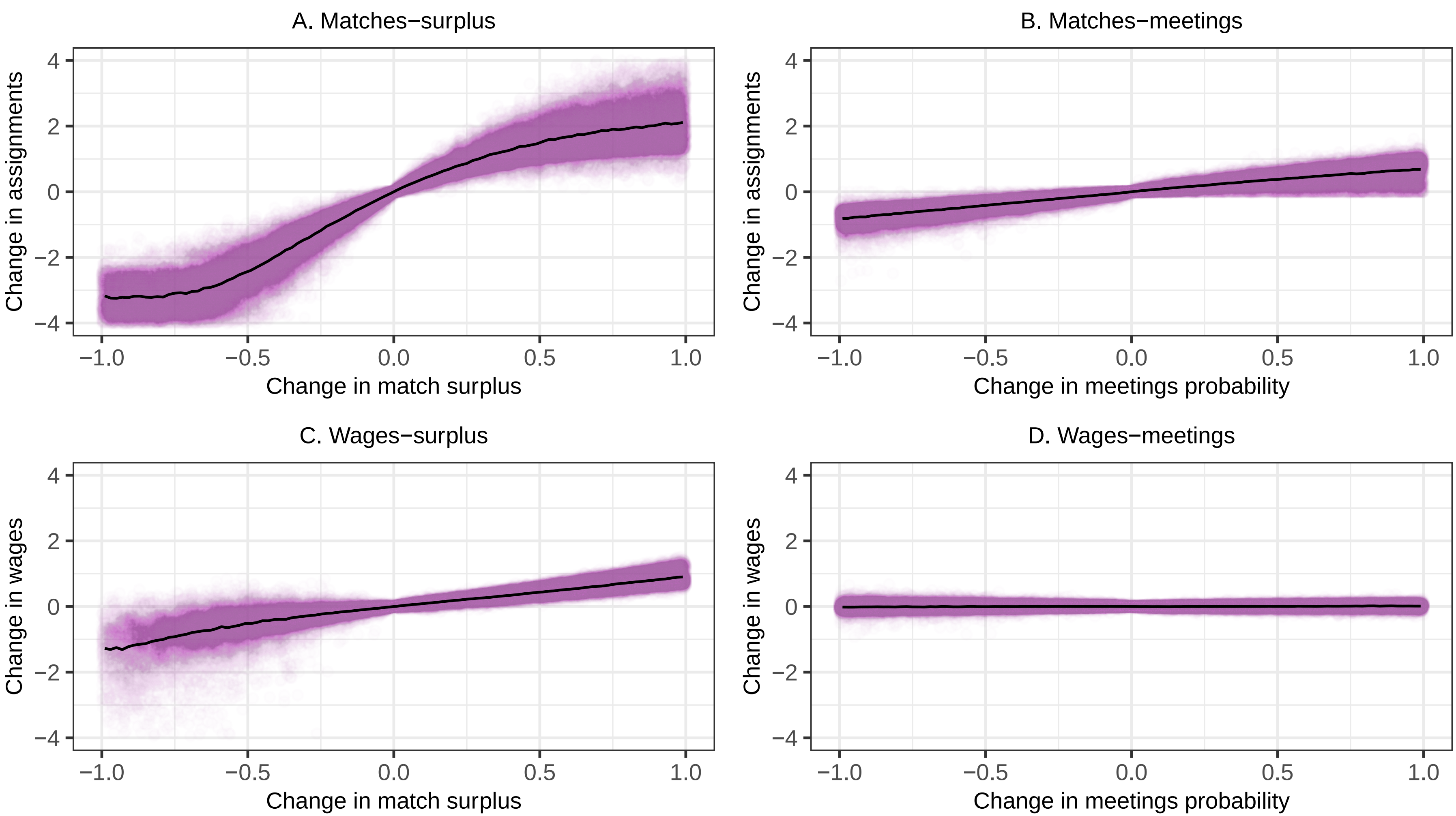
Parameters and moments

- Parameters
  - Meeting probabilities: p_{txyc}
  - Systematic match utility: β_{txyc}
  - Idiosyncratic utility scale: σ
  - (Utility location: b)
- Moments
  - Number of matches: μ_{txyc}
  - Average wage: w_{txyc}
  - Within-group wage variance: $WithinVar_w$
  - (Wage variance: Var_w)

Groups and observations

- $T = 10$ (2006-2015)
- $X = 8$ (Jews/Arabs $\times$ no-college/college $\times$ males/females)
- $Y = 5$ (bins of AKM firm premiums)
- $C = 4$ (none, phantom, weak, and strong)
- $I \approx 200K$

Identification of the model



elasticities table

Estimation: inverting the data (outer loop)

- Use an update mapping that "inverts" the data into the parameters

$$
\begin{aligned}
p_n^{h+1} &= p_n^h + \eta \left[log(\mu_n) - log(\hat{\mu}_n(p^h, \beta^h)) \right] \\
\beta_n^{h+1} &= \beta_n^h + \eta \left[log(\mu_n \cdot w_n) - log(\hat{\mu}_n(p^h, \beta^h) \cdot \hat{w}_n(p^h, \beta^h)) \right]
\end{aligned}
$$

where

- Parameters:
  - p: meeting rate
  - β: match utility
- Moments:
  - μ: matches share
  - w: average wage
- h: iteration index
- $n \equiv txyc$: a combination of market t, worker group x, firm group y, and connection type c
- $\eta > 0$: update rate

full update mapping

Model fit

Table 4: Model's fit and precision

A. Model's fit

	Matches (μ_{txyc}) (1)	Av. wage (w_{txyc}) (2)	Overall wage variance (3)	Within-group wage variance (4)
Abs. deviation	0.013	0.008	0.0008	0.0007
	(0.0006)	(0.0006)	(0.0006)	(0.0005)
Correlation	1.000	0.998		
	(0.00002)	(0.0002)		

B. Model's precision and Monte Carlo simulation

	Surplus (β_{txyc}) (1)	Meetings (p_{txyc}) (2)	Unobserved heterogeneity ($log(\sigma)$) (3)	Surplus scale (b) (4)
Estimates				
Correlation	0.980	0.988		
	(0.001)	(0.0006)		
Value			-1.069	9.174
			(0.007)	(0.011)
Monte Carlo				
Correlation	0.972	0.985		
	(0.003)	(0.0006)		
Value			-1.076	9.186
			(0.006)	(0.009)

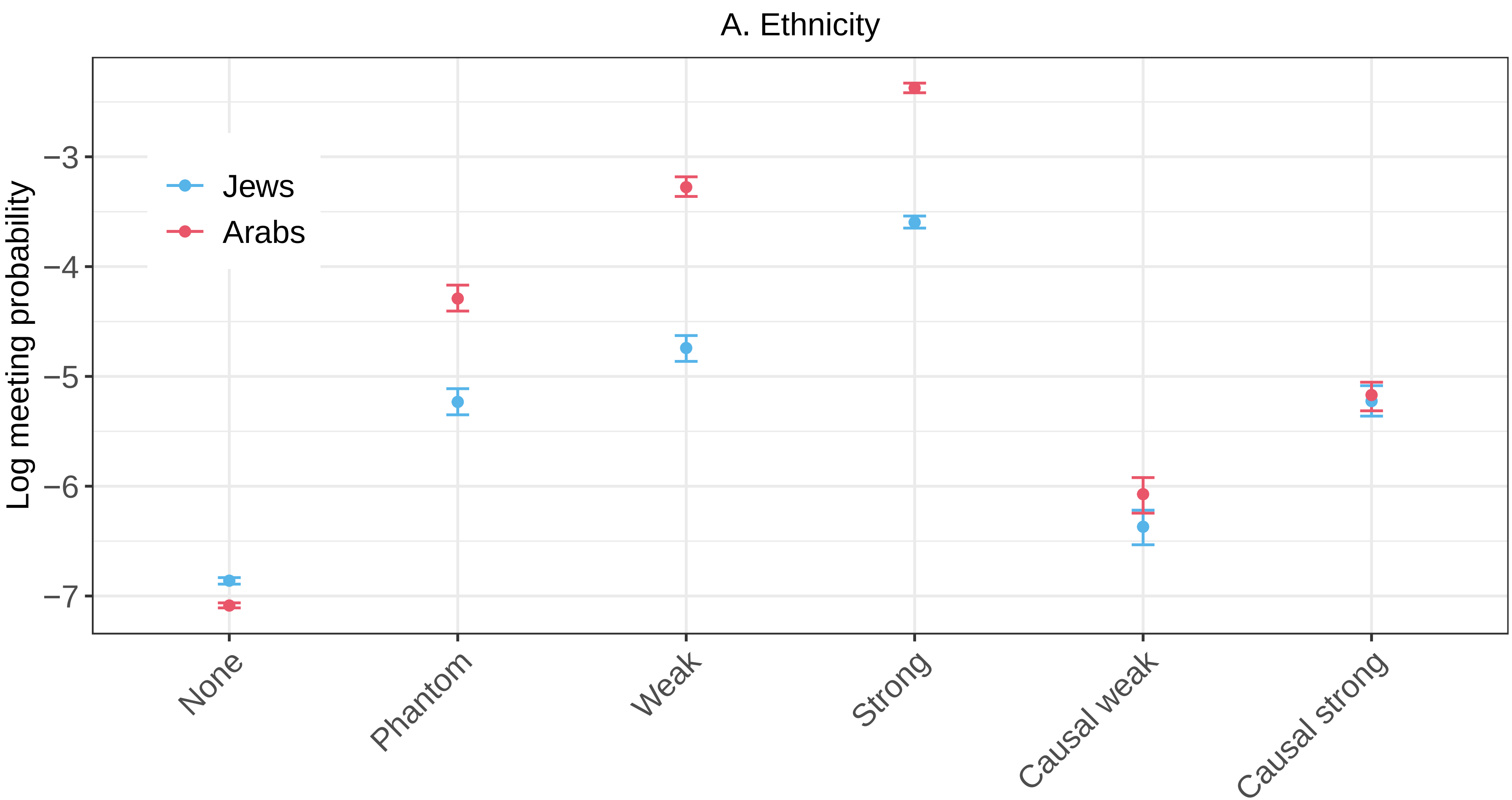
Outline

Model estimates

Table 5: Projection of the model estimates on workers', firms', and connections' characteristics

	Meeting probability ($Log(p_{txyc})$)	Firm's surplus (β_{txyc})
	(1)	(2)
Constant	-6.900	8.809
	(0.015)	(0.011)
Phantom connections	1.964	0.012
	(0.039)	(0.007)
Weak connections	2.728	0.041
	(0.038)	(0.008)
Strong connections	3.742	0.158
	(0.019)	(0.004)
Arab	0.051	-0.011
	(0.010)	(0.002)
Female	-0.009	-0.070
	(0.010)	(0.002)
College	-0.066	0.077
	(0.011)	(0.002)
Job type: 2	-0.067	0.120
	(0.012)	(0.005)
Job type: 3	-0.028	0.268
	(0.012)	(0.005)
Job type: 4	-0.002	0.459
	(0.013)	(0.006)
Job type: 5	-0.093	0.967
	(0.021)	(0.007)
Weak - phantom	0.764	0.028
	(0.054)	(0.010)
Strong - phantom	1.779	0.146
	(0.042)	(0.008)
R^2	0.831	0.907
	(0.005)	(0.003)

Meeting probability by ethnicity and connections type



by gender by bargaining power

Outline

Value of a meeting

Table 6: Value of meetings and connections

	Total expected gains	Salary change with a job change			Salary change without a job change		
		Probability	Gains	Expected gains	Probability	Gains	Expected gains
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
New meeting, without surplus effect	2.2	0.040	41.4	1.7	0.064	7.9	0.5
	(0.417)	(0.007)	(6.543)	(0.394)	(0.008)	(1.809)	(0.135)
Existing meeting, with surplus effect	1.5	0.040	20.3	0.8	0.101	6.4	0.7
	(0.467)	(0.007)	(8.151)	(0.373)	(0.010)	(2.974)	(0.311)
New meeting, with surplus effect	3.7	0.055	57.0	3.1	0.066	9.0	0.6
	(0.819)	(0.009)	(9.323)	(0.778)	(0.008)	(2.248)	(0.153)

by job type

Between-group pay gaps

Table 7: Counterfactual impacts of connections on between-group pay gaps

A. Equalizing number of connections per worker

	Gap	Without identification strategy			With identification strategy		
	(% Average)	Meetings effect	Surplus effect	Both effects	Meetings effect	Surplus effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity gap	-8.4	-59.5	-0.4	-67.6	-5.1	-1.1	-11.7
	(0.351)	(4.866)	(0.168)	(3.031)	(0.679)	(0.297)	(1.638)
Gender gap	-18.0	1.2	0.0	2.3	0.1	0.0	0.1
	(0.290)	(0.180)	(0.034)	(0.197)	(0.066)	(0.045)	(0.093)

B. Prohibiting hiring of connected workers

	Baseline	Weak	Strong	Weak + strong
	(% Average)			
	(1)	(2)	(3)	(4)
Ethnicity gap	-8.4	8.9	44.3	56.4
	(0.351)	(0.982)	(2.820)	(3.347)
Gender gap	-18.0	-4.0	-20.3	-25.3
	(0.290)	(0.320)	(0.780)	(0.798)

pay-premium utility efficiency

Outline

Review

- In Israel, (weak) parental connections increase hiring in a firm by
  - 3.7 times (regression)
  - 2.9 times (model)
    - 115% search frictions + 35% match value
  - Stronger effect for Arabs
- Value of one additional meeting with a connected firm is 3.7% the average wage
  - 2.2% search frictions + 1.5% match value
  - 3.1% direct (changing job) + 0.6% indirect (better choice set)
- Impacts of connections on ethnic pay gaps
  - Equalizing connections: pay gap decreases by 12%
    - 5% without the match-value effect
  - Prohibiting connections: pay gap increases by 56%

Thank you!

References I

Abowd, John M., Francis Kramarz, and David N. Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Athey, Susan, Christopher Avery, and Peter Zemsky, "Mentoring and diversity," *American Economic Review*, 2000, *90* (4), 765–786.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social connections and incentives in the workplace: Evidence from personnel data," *Econometrica*, 2009, *77* (4), 1047–1094.

Beaman, Lori and Jeremy Magruder, "Who gets the job referral? Evidence from a social networks experiment," *American Economic Review*, 2012, *102* (7), 3574–93.

References II

Bernard, Florian, Nikos Vlassis, Peter Gemmar, Andreas Husch, Johan Thunberg, Jorge Goncalves, and Frank Hertel, "Fast correspondences for statistical shape models of brain structures," in "Medical Imaging 2016: Image Processing," Vol. 9784 International Society for Optics and Photonics 2016, p. 97840R.

Bertsekas, Dimitri P., *Network optimization: continuous and discrete models*, Athena Scientific Belmont, MA, 1998.

Bolte, Lukas, Nicole Immorlica, and Matthew O. Jackson, "The Role of Referrals in Inequality, Immobility, and Inefficiency in Labor Markets," *Unpublished*, 2020.

Bonnet, Odran, Alfred Galichon, Keith O'Hara, and Matthew Shum, "Yogurts Choose Consumers? Estimation of Random Utility Models via Two-Sided Matching," *Unpublished*, 2018.

Caldwell, Sydnee and Nikolaj Harmon, "Outside Options, Bargaining, and Wages: Evidence from Coworker Networks," *Unpublished*, 2018, p. 107.

Calvo-Armengol, Antoni and Matthew O. Jackson, "The effects of social networks on employment and inequality," *American economic review*, 2004, *94* (3), 426–454.

Card, David, Ana Rute Cardoso, Jörg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

Choo, Eugene and Aloysius Siow, "Who marries whom and why," *Journal of political Economy*, 2006, *114* (1), 175–201.

Cingano, Federico and Alfonso Rosolia, "People I know: job search and social networks," *Journal of Labor Economics*, 2012, *30* (2), 291–332.

Corak, Miles and Patrizio Piraino, "The intergenerational transmission of employers," *Journal of Labor Economics*, 2011, *29* (1), 37–68.

References IV

Demange, Gabrielle and David Gale, "The strategy structure of two-sided matching markets," *Econometrica: Journal of the Econometric Society*, 1985, pp. 873–888. Publisher: JSTOR.

Dickinson, David L., David Masclet, and Emmanuel Peterle, "Discrimination as favoritism: The private benefits and social costs of in-group favoritism in an experimental labor market," *European Economic Review*, 2018, *104*, 220–236.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-based job search networks," *The Review of Economic Studies*, 2016, *83* (2), 514–546.

Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordstrom Skans, "Social Connections and the Sorting of Workers to Firms," *Unpublished*, 2019.

Fontaine, Francois, "Why are similar workers paid differently? The role of social networks," *Journal of Economic Dynamics and Control*, 2008, *32* (12), 3960–3977. Publisher: Elsevier.

Galichon, Alfred and Bernard Salanié, "Cupid's invisible hand: Social surplus and identification in matching models," *Unpublished*, 2015.

Granovetter, Mark, *Getting a job: A study of contacts and careers*, University of Chicago press, 1973.

Kramarz, Francis and Oskar Nordström Skans, "When strong ties are strong: Networks and youth labour market entry," *Review of Economic Studies*, 2014, *81* (3), 1164–1200.

Montgomery, James D., "Social networks and labor-market outcomes: Toward an economic analysis," *The American economic review*, 1991, *81* (5), 1408–1418.

Plug, Erik, Bas van der Klaauw, and Lennart Ziegler, "Do Parental Networks Pay Off? Linking Children's Labor-Market Outcomes to Their Parents' Friends," *The Scandinavian Journal of Economics*, 2018, *120* (1), 268–295.

Shapley, Lloyd S. and Martin Shubik, "The assignment game I: The core," *International Journal of game theory*, 1971, *1* (1), 111–130.

Topa, Giorgio, "Labor markets and referrals," in "Handbook of social economics," Vol. 1, Elsevier, 2011, pp. 1193–1221.

Sample selection

- **Full sample**: panel dataset at the annual frequency
  - Ages 22-80
  - Assigning the firm with the maximal salary in February
  - Excluding worker-year observations $<$ 25% the national average monthly wage
- **5-500 sample**: firms with 5-500 workers
- **New workers sample**: the first real job of workers
  - Natives, ages 22-27 at 2006-2015
  - First job after graduation, 5-500 firm, $\geq$ 4 months, annual earnings $\geq$ 150% the national average monthly wage (Kramarz and Skans 2014)
  - Graduation year $=$ 21 for workers with no college

back

Parental connections

- Three types of connections between a new worker i and firm j
  - Weak connections
    - i's parent and k worked simultaneously at $j' \neq j$ when i was 12-21 years old
    - k worked at j at time 0 (= the year i entered the labor market)
  - Phantom connections
    - i's parent and k worked simultaneously at $j' \neq j$ when i was 12-21 years old
    - k worked at j at time [-5,5] but not at time 0
  - Strong connections
    - i's parent worked at j when i was 12-21 years old, or
    - i has at least two weak or phantom contacts at j
- All firms belong to the 5-500 sample

back

Firm pay premium

- Estimating AKM model (Abowd et al. 1999)

$$w_{it} = \alpha_i + \psi_{J(it)} + Z'_{it}\gamma + \varepsilon_{it}$$

with

- α_i = person FE
- $\psi_{J(it)}$ = firm FE
- Z'_{it} = year FEs, and quartic polynomials of age restricted to be flat at age 40 (Card et al. 2018)
- Firm premium at year t is calculated using the largest connected set of the full sample at years [t-4,t]
- Firms are ranked within year

back

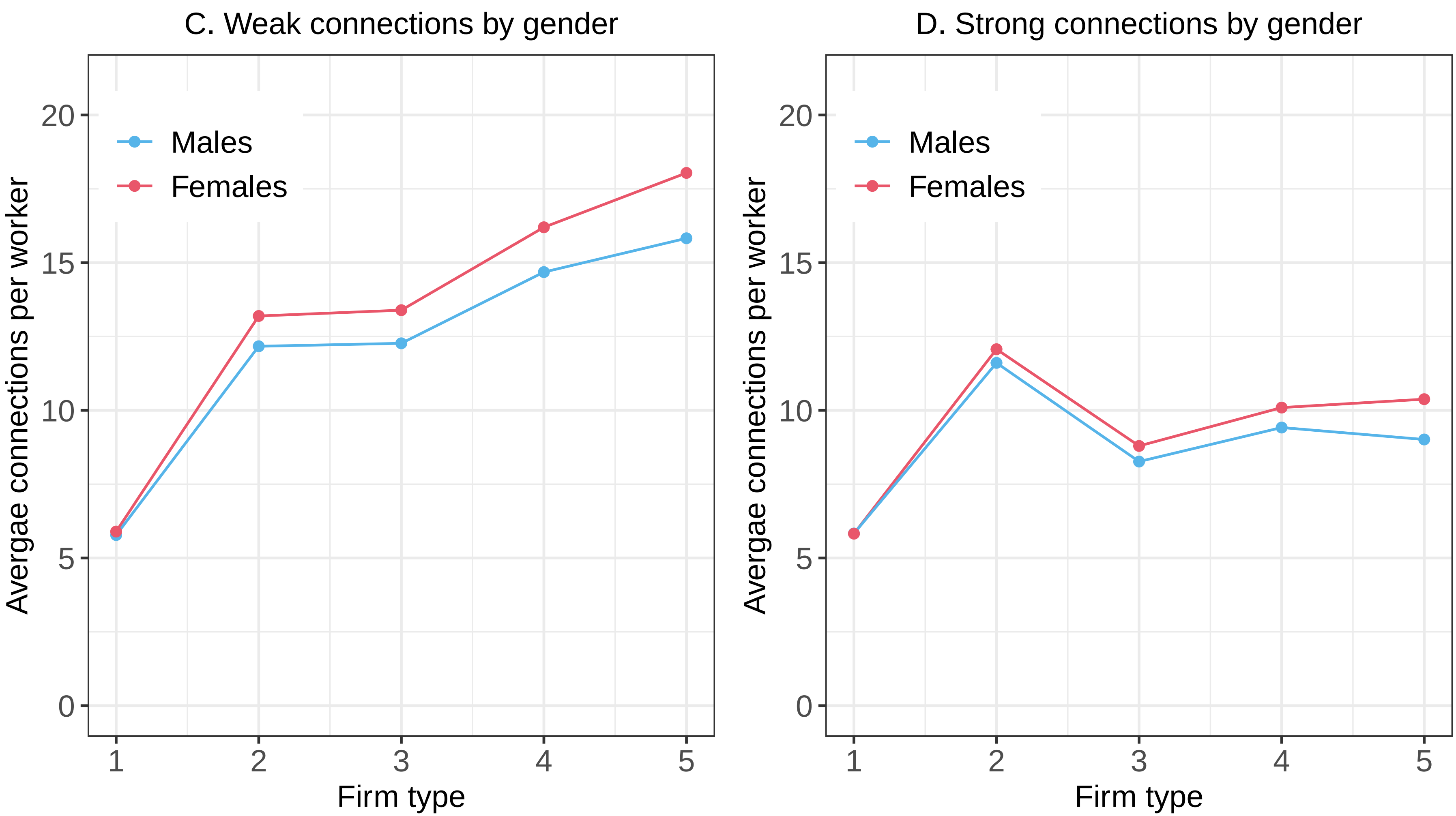
Raw ethnic and gender pay gaps

Table 8: Earnings gap by ethnicity and gender, new workers

	Log salary			
	(1)	(2)	(3)	(4)
Arab	-0.077	0.030	-0.062	0.030
	(0.004)	(0.003)	(0.004)	(0.003)
Female	-0.203	-0.134	-0.203	-0.134
	(0.003)	(0.002)	(0.003)	(0.002)
Weak con qualiy			0.117	-0.001
			(0.010)	(0.008)
Strong con qualiy			0.090	-0.014
			(0.007)	(0.006)
Firm FE	No	Yes	No	Yes
Observations	211,144	211,144	211,144	211,144
N firms	52,963	52,963	52,963	52,963
R^2 (full model)	0.138	0.614	0.140	0.614
R^2 (projected model)	0.080	0.047	0.083	0.047

back

Connections per worker by gender



back

Balancing test

Table 9: Balancing test: Correlation between parental connections and measures of proximity between workers and firms

	Log distance	Parent's industry
	(1)	(2)
Phantom connections	-0.369	0.077
	[-0.376,-0.362]	[0.076,0.077]
Weak connections	-0.368	0.076
	[-0.375,-0.361]	[0.075,0.076]
Strong connections	-0.926	0.281
	[-0.944,-0.909]	[0.279,0.284]
R0 (no connections)	10.102	0.033
	[10.090,10.117]	[0.032,0.033]
Ratio weak-phantom	1.000	0.989
	[1.000,1.001]	[0.984,0.995]
Ratio strong-phantom	0.943	2.871
	[0.942,0.944]	[2.850,2.887]
Observations (firms x groups)	21,166,443	21,166,443
N firms	149,729	149,729
N groups	2,959	2,959
N workers	220,684	220,684

back

Exogenous separations

- Use death and retirement of contacts for exogenous separation causes

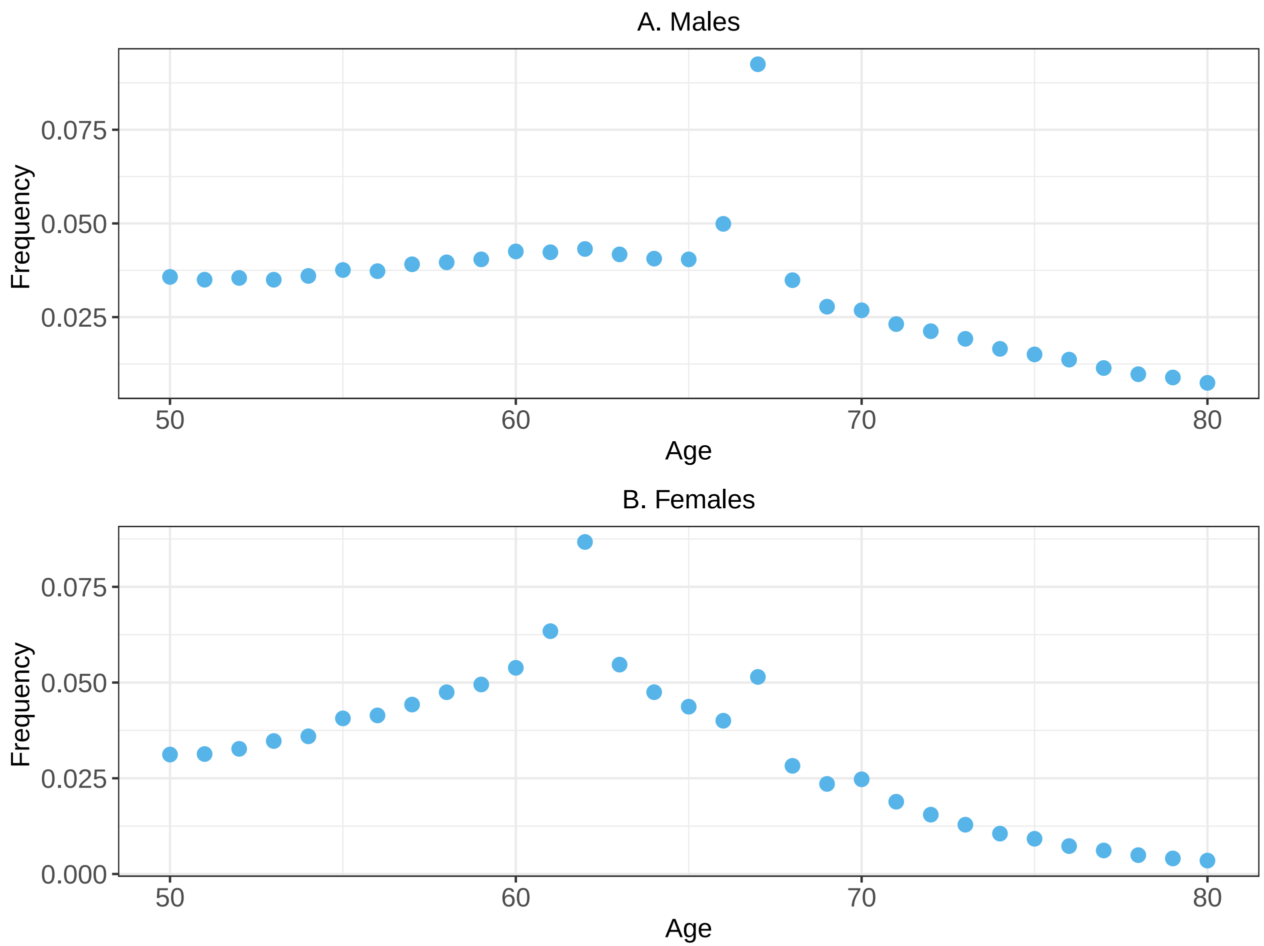
Death and retirement of contacts

Table 10: Effects of parental connections on firm assignment: death and retirement of contacts

	Employment		
	(1)	(2)	(3)
Special connections:	Death	Retirement	Death or retirement
Phantom (D/R)	0.031	0.010	0.017
	[0.004,0.068]	[-0.008,0.032]	[0.001,0.034]
Phantom (Other)	0.010	0.010	0.010
	[0.009,0.011]	[0.009,0.011]	[0.009,0.011]
Weak (D/R)	0.065	0.032	0.041
	[0.010,0.126]	[0.003,0.066]	[0.017,0.071]
Weak (Other)	0.050	0.051	0.051
	[0.047,0.054]	[0.047,0.055]	[0.047,0.054]
Strong	0.487	0.487	0.487
	[0.472,0.501]	[0.472,0.501]	[0.472,0.501]
R0 (no connections)	0.005	0.005	0.005
	[0.005,0.005]	[0.005,0.005]	[0.005,0.005]
Ratio weak-phantom (D/R)	2.567	3.913	2.773
	[0.386,7.746]	[0.582,19.460]	[0.748,6.533]
Ratio weak-phantom (Other)	3.679	3.680	3.691
	[3.335,4.101]	[3.339,4.099]	[3.349,4.122]
N connections: phantom (D/R)	85,532	138,194	222,461
N connections: weak (D/R)	37,402	102,499	138,974

age at retirement back

Age at retirement

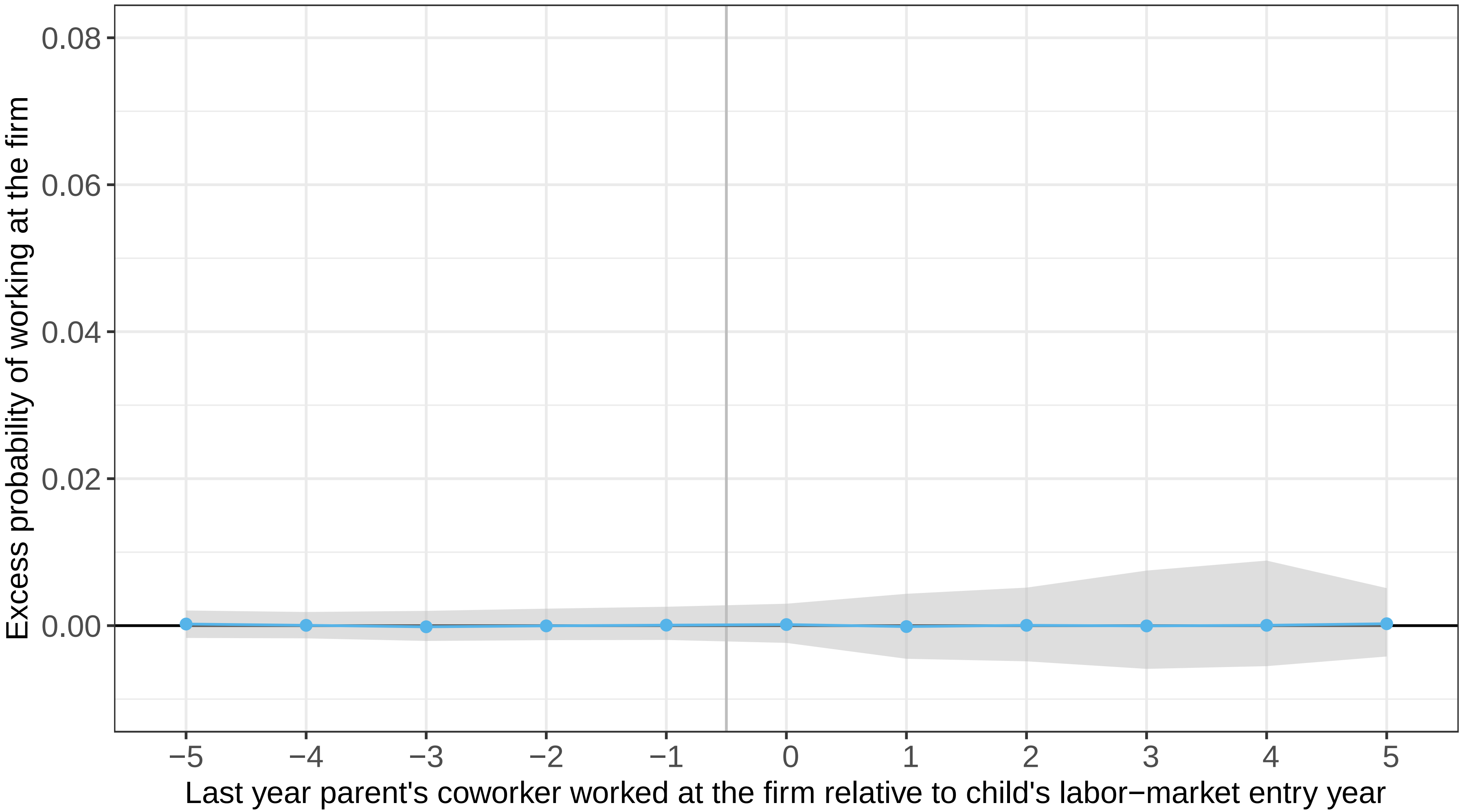


back

Placebo test

- Assigning to each worker the connections of a random worker in her group

Placebo test: event study



Placebo test: Average effects

Table 11: Effect of weak parental connections on firm assignment, placebo test

	All	Jews	Arabs	Males	Females
	(1)	(2)	(3)	(4)	(5)
Phantom connections	0.000	0.000	0.000	0.000	0.000
	[-0.001,0.001]	[-0.001,0.001]	[-0.002,0.003]	[-0.001,0.001]	[-0.001,0.001]
Weak connections	0.000	0.000	0.000	0.000	0.000
	[-0.002,0.002]	[-0.002,0.002]	[-0.006,0.006]	[-0.002,0.003]	[-0.003,0.003]
Strong connections	0.000	0.000	0.001	0.000	0.000
	[-0.006,0.007]	[-0.005,0.005]	[-0.021,0.021]	[-0.006,0.008]	[-0.008,0.010]
R0 (no connections)	0.007	0.006	0.011	0.008	0.007
	[0.007,0.008]	[0.006,0.007]	[0.011,0.012]	[0.007,0.008]	[0.007,0.007]
Ratio weak-phantom	1.010	1.000	1.053	1.011	1.017
	[0.755,1.384]	[0.727,1.330]	[0.397,1.645]	[0.660,1.334]	[0.631,1.524]
Ratio strong-phantom	1.047	1.029	1.107	1.065	1.036
	[0.206,2.019]	[0.189,1.805]	[-0.938,3.233]	[0.154,1.981]	[-0.162,2.471]
Observations	21,166,443	16,837,526	4,328,917	15,319,313	5,847,130
N firms	149,729	144,186	117,746	145,939	134,555
N groups	2,959	1,658	1,301	1,548	1,411
N workers	220,684	157,009	63,675	170,872	49,812
N connections	40,827,833	33,261,814	7,566,019	31,664,340	9,163,493

back

Robustness checks: definitions of connections

Table 12: Effects of parental connections on firm assignment: Robustness to the definition of connection types

	Employment		
	(1)	(2)	(3)
Phantom (single contact)	0.010	0.012	
	[0.009,0.011]	[0.011,0.013]	
Phantom (single + multiple contacts)			0.015
			[0.014,0.016]
Weak (signle contact)	0.050	0.053	
	[0.047,0.054]	[0.049,0.056]	
Weak (single + multiple contacts)			0.095
			[0.091,0.100]
Strong (direct + multiple contacts)	0.487		
	[0.472,0.501]		
Direct		3.091	3.092
		[2.977,3.206]	[2.978,3.207]
Multiple contacts		0.171	
		[0.161,0.181]	
R0 (no connections)	0.005	0.005	0.005
	[0.005,0.005]	[0.005,0.005]	[0.005,0.005]
Observations (firms x groups)	21,166,443	21,166,443	21,166,443
N firms	149,729	149,729	149,729
N groups	2,959	2,959	2,959
N workers	220,684	220,684	220,684
N connections	40,827,833	40,827,833	40,827,833

back

Heterogeneity: stylized facts

- Connections are stronger if generated
  - In smaller firms
  - In longer periods
  - More recently
  - Between similar individuals
- The effect is stronger for
  - Males
  - Arabs
  - No-college workers

back

Equilibrium

- An equilibrium outcome (μ, w) consist of an equilibrium matching $\mu(i,j)$ and an equilibrium wage $w(i,j)$ such that:
  1. Matching $\mu(i,j)$ is feasible:

$$\sum_j \mu(i,j) \leq 1 \quad , \quad \sum_i \mu(i,j) \leq 1 \quad , \quad \mu(i,j) = 1 \implies m(i,j) = 1$$

  2. Matching $\mu(i,j)$ is optimal for workers and firms given wages w and meetings m:

$$\mu(i,j) = 1 \implies j \in \text{argmax}_{j \in m_i} U_{ij} \quad \text{and} \quad i \in \text{argmax}_{i \in m_j} V_{ij}$$

back

Auction algorithm I

1. Start with an empty assignment S, a vector of initial wages w_i, and some $\epsilon > 0$
2. Iterate on the two following phases:
    1. Bidding Phase
    For each unassigned firm j in the assignment S :
        1. Find a "best" worker $i_j \in m(j)$ having maximum value and the corresponding value

$$i_j = \arg\max_{i \in m(j)} f_{ij} - w_i \quad , \quad v_j = \max_{i \in m(j)} f_{ij} - w_i$$

and find the best value offered by workers other than i_j

$$q_j = \max_{i \in m(j), i \neq i_j} f_{ij} - w_i$$

Auction algorithm II

2. Compute the "bid" of firm j given by

$$b_{ij} = w_{i_j} + v_j - q_j + \epsilon$$

2. Assignment Phase
   For each worker i, let $B(i)$ be the set of firms from which i received a bid. If $B(i)$ is non-empty, increase w_i to the highest bid:

$$w_i = \max_{j \in B(i)} b_{ij} \tag{1}$$

   and assign i to the firm in $B(i)$ attaining the maximum above

3. Terminate when all workers are assigned to firms

back

Bellman-Ford algorithm

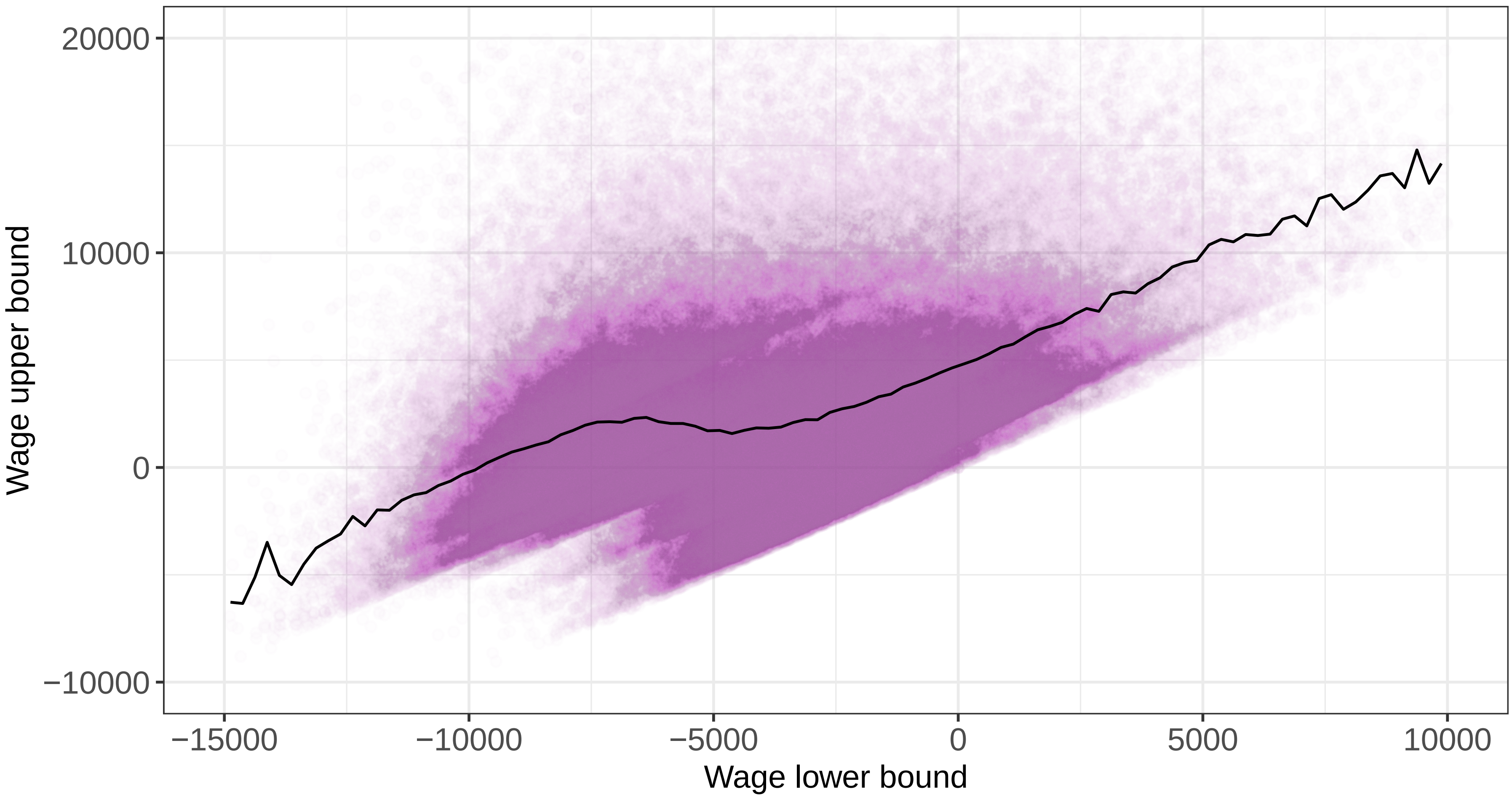
- The firm-optimal equilibrium wages are the fixed point of the mapping

$$w_i = \max(w_i, \max_{j \in m(i)} (f_{ij} - v_j)) \ , \ v_j = \min(v_j, f_{i*(j)j} - w_{i*(j)}) \ , \ w_0 = 0$$

  - $i^*(j)$ denote the equilibrium match of firm j
- The fixed point can be computed by iterating on the map from the initial values $\{w_i = -\infty, w_0 = 0; v_j = \infty\}$
- The worker-optimal equilibrium wages can be found similarly
- The bounds are finite iff each connected set is a double connected set

back

Lower and upper wage bounds



back

Simulating an equilibrium outcome (inner loop)

- Given parameters and a draw of unobservables:
  1. Get the set of meetings m_{ij}
  2. Calculate the joint surplus f_{ij}
  3. Find the equilibrium matching using the auction algorithm
  4. Find the equilibrium wage using the BF algorithm
- The two-stage model offers a computational advantage over existing matching models
- Exploit the sparsity of the data using c++ implementations of the auction (Bernard et al. 2016) and BF algorithms

back

Moments-parameters elasticities

Table 13: Moments-parameters elasticities

	Matches-surplus	Matches-meetings	Wages-surplus	Wages-meetings
	$dln(\mu)/d\beta$	$dln(\mu)/dln(p)$	$dln(w)/d\beta$	$dln(w)/dln(p)$
	(1)	(2)	(3)	(4)
Same workers and firms	3.511	0.777	3.427	0.015
	(0.078)	(0.017)	(0.325)	(0.009)
Same workers, different firms	-0.264	-0.033	0.001	0.014
	(0.026)	(0.003)	(0.011)	(0.001)
Different workers	-0.008	0.000	-0.032	-0.002
	(0.002)	(0.000)	(0.005)	(0.000)

back

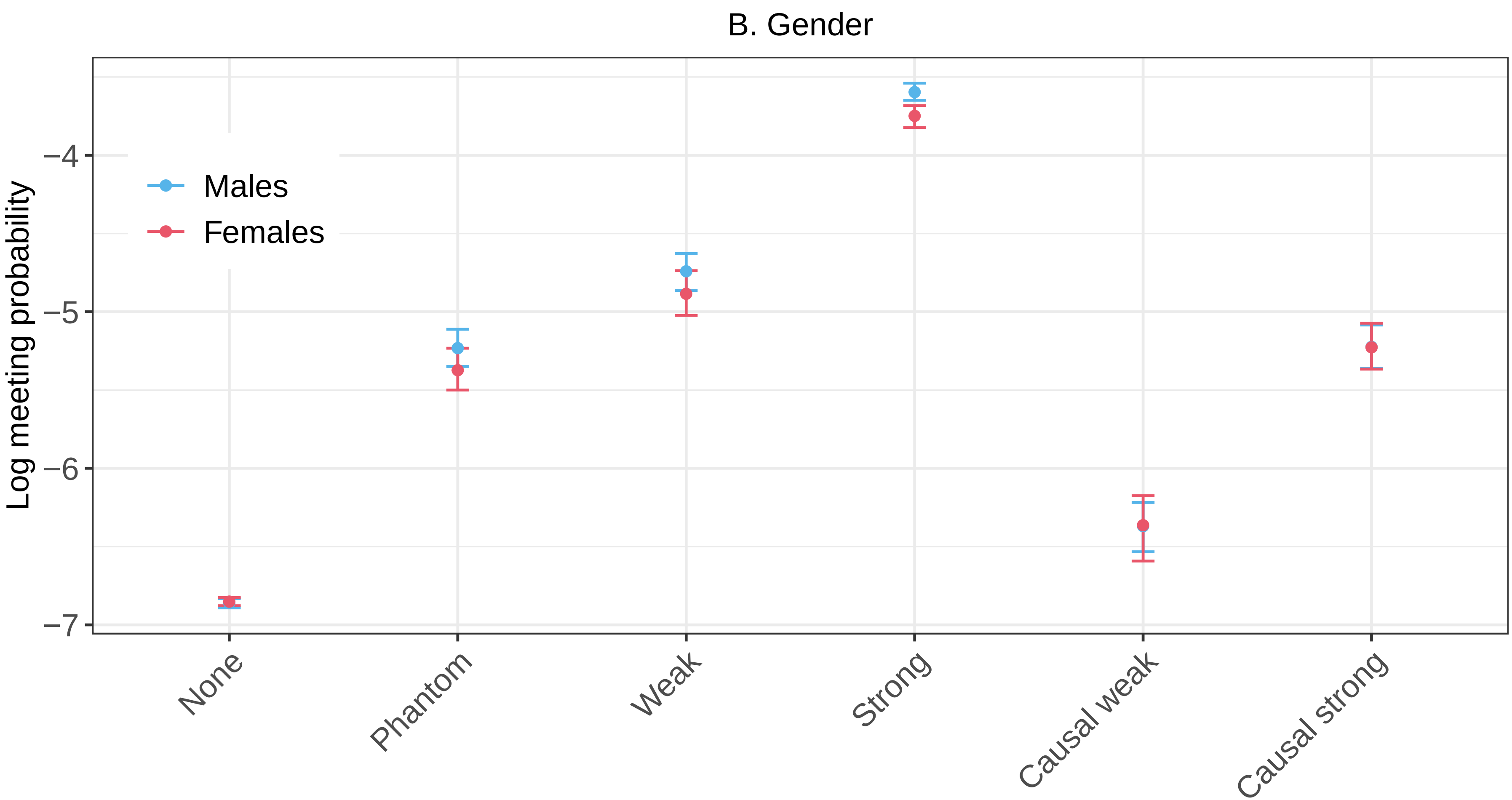
Estimation: inverting the data (outer loop)

$$p_n^{h+1} = p_n^h + \eta \left[log(\mu_n) - log(\hat{\mu}_n(p^h, \beta^h, \sigma^h, b^h)) \right]$$
$$\beta_n^{h+1} = \beta_n^h + \eta \left[log(\mu_n \cdot w_n) - log(\hat{\mu}_n(p^h, \beta^h, \sigma^h, b^h) \cdot \hat{w}_n(p^h, \beta^h, \sigma^h, b^h)) \right]$$
$$\sigma^{h+1} = \sigma^h + \eta \left[log(WithinVar_w) - log(Wit\hat{hin}Var_w(p^h, \beta^h, \sigma^h, b^h)) \right]$$
$$b^{h+1} = b^h + \eta \left[log(Var_w) - log(\hat{Var}_w(p^h, \beta^h, \sigma^h, b^h)) \right]$$

where

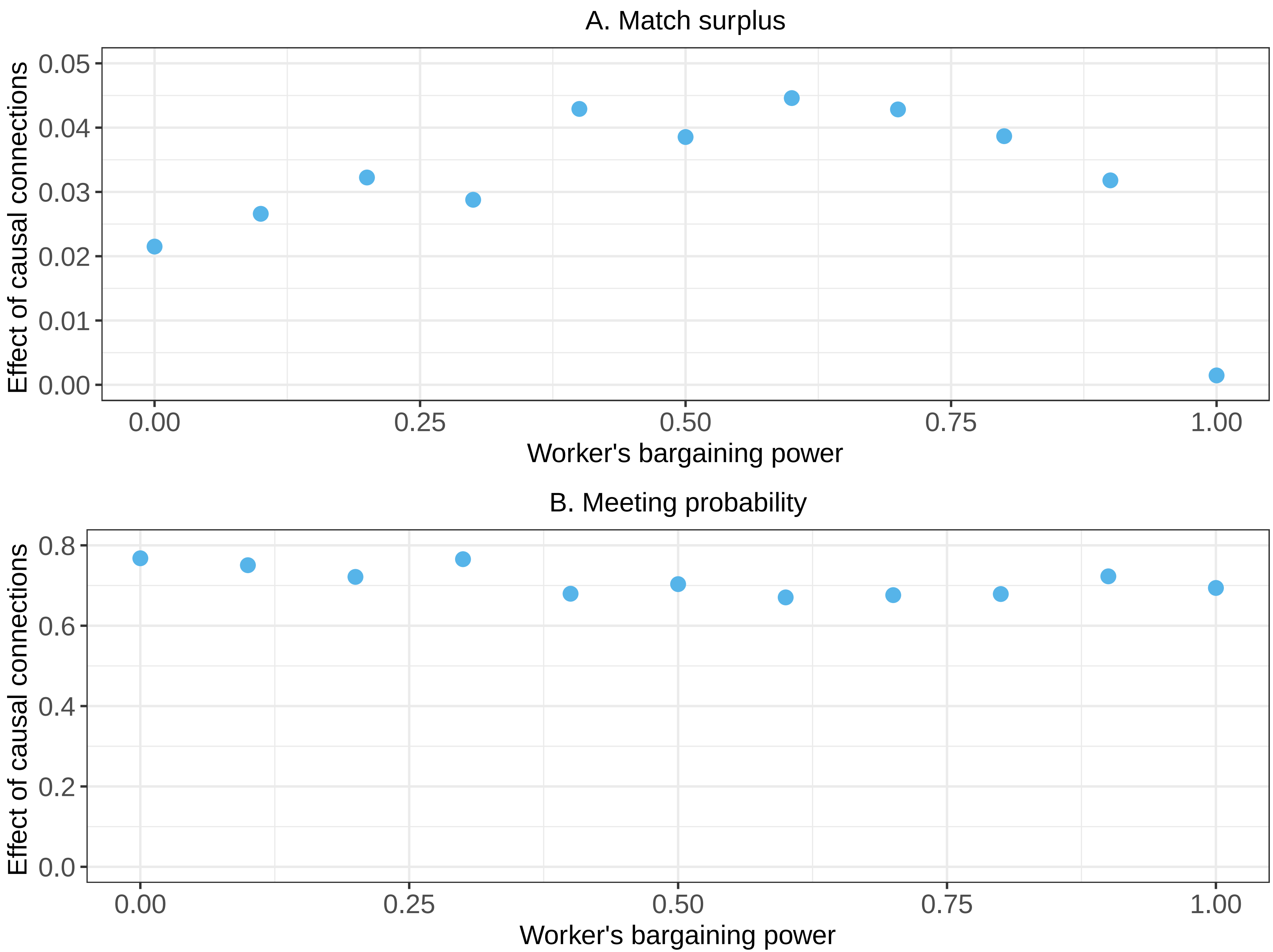
- Parameters:
  - p: meeting rate; β: match utility; σ: idiosyncratic utility scale; b: utility location
- Moments:
  - μ: matches share; w: average wage; Var_w: overall wage variance; $WithinVar_w$: within-group wage variance
- $n \equiv txyc$: a combination of market t, worker group x, firm group y, and connection type c
- $\eta > 0$: update rate

Meeting probability by gender and connections type

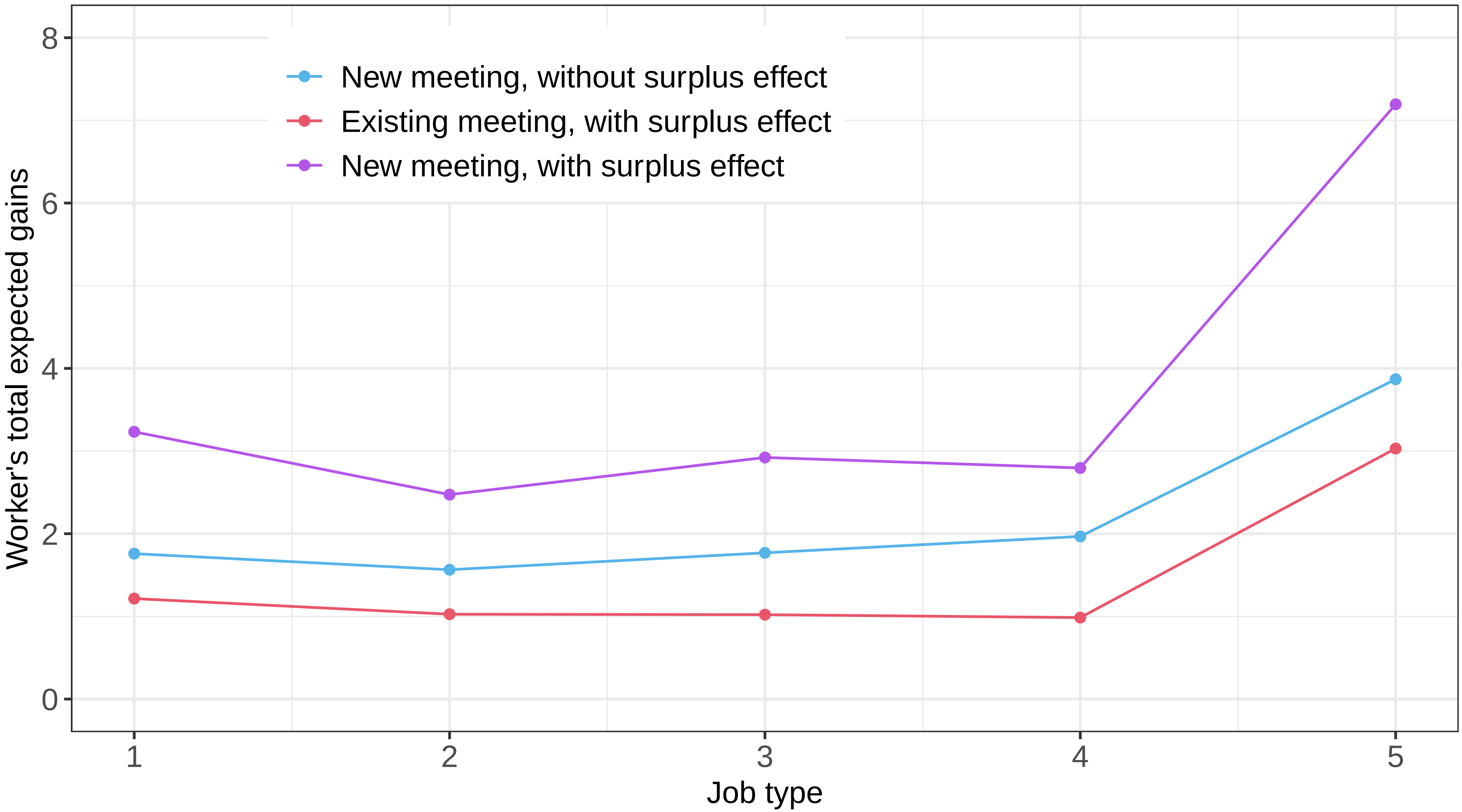


back

Model estimates by worker's bargaining power



Value of a meeting/connection by job type



back

Between-group pay-premium gaps

Table 14: Counterfactual impacts of connections on between-group gaps in firm pay premiums

A. Equalizing number of connections per worker

	Gap	Without identification strategy			With identification strategy		
	(% Average)	Meetings effect	Surplus effect	Both effects	Meetings effect	Surplus effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity gap	-23.1	-15.3	-0.1	-15.2	-1.4	-0.1	-2.4
	(0.299)	(1.500)	(0.180)	(0.754)	(0.326)	(0.204)	(0.502)
Gender gap	2.1	0.0	0.1	1.2	0.5	0.1	1.4
	(0.268)	(3.318)	(1.412)	(3.479)	(1.794)	(1.560)	(2.402)

B. Prohibiting hiring of connected workers

	Baseline	Weak	Strong	Weak + strong
	(% Average)			
	(1)	(2)	(3)	(4)
Ethnicity gap	-23.1	-0.9	-1.6	-2.8
	(0.299)	(0.511)	(0.835)	(0.955)
Gender gap	2.1	8.0	36.3	46.2
	(0.268)	(4.775)	(11.271)	(11.609)

back

Between-group utility gaps

Table 15: Counterfactual impacts of connections on between-group gaps in match utility

A. Equalizing number of connections per worker

	Gap	Without identification strategy			With identification strategy		
	(% Average)	Meetings effect	Surplus effect	Both effects	Meetings effect	Surplus effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity gap	-17.8	-20.8	-0.2	-21.6	-1.8	-0.3	-3.8
	(0.297)	(2.053)	(0.168)	(0.944)	(0.372)	(0.205)	(0.700)
Gender gap	-6.8	1.1	0.0	1.9	-0.1	0.0	-0.2
	(0.310)	(0.705)	(0.274)	(0.755)	(0.365)	(0.334)	(0.485)

B. Prohibiting hiring of connected workers

	Baseline	Weak	Strong	Weak + strong
	(% Average)			
	(1)	(2)	(3)	(4)
Ethnicity gap	-17.8	0.3	4.1	4.6
	(0.297)	(0.436)	(0.808)	(0.850)
Gender gap	-6.8	-5.1	-27.5	-33.9
	(0.310)	(1.016)	(2.102)	(2.232)

back

Impacts on overall efficiency

Table 16: Counterfactual impacts of connections on efficiency

A. Equalizing number of connections per worker

	Without identification strategy			With identification strategy		
	Meetings effect	Surplus effect	Both effects	Meetings effect	Surplus effect	Both effects
	(1)	(2)	(3)	(4)	(5)	(6)
Equilizing connections by Ethnicity	0.4	0.0	0.5	0.0	0.0	0.1
	(0.032)	(0.001)	(0.015)	(0.005)	(0.003)	(0.014)
Equilizing connections by gender	0.1	0.0	0.1	0.0	0.0	0.0
	(0.005)	(0.001)	(0.005)	(0.002)	(0.001)	(0.003)

B. Prohibiting hiring of connected workers

	Weak	Strong	Weak + strong
	(1)	(2)	(3)
Prohibiting connected hiring	-0.4	-2.2	-2.6
	(0.011)	(0.026)	(0.030)

back