The Peer Effect on Future Wages in the Workplace

Long Hong[1] Salvatore Lattanzio[2]

[1]University of Wisconsin - Madison

[2]Bank of Italy

EEA-ESEM 2022, Bocconi University, Milan

August 22-26, 2022

Motivation

Workplace interaction is a critical component of human capital accumulation

Motivation

Workplace interaction is a critical component of human capital accumulation

- Network effects
- Peer pressure
- Learning spillovers

Motivation

Workplace interaction is a critical component of human capital accumulation

- Network effects
- Peer pressure
- Learning spillovers

Sorting among coworkers may explain how wages grow differently for different workers

- Little evidence that quantifies empirically this contribution

Motivation

Workplace interaction is a critical component of human capital accumulation

- Network effects
- Peer pressure
- Learning spillovers

Sorting among coworkers may explain how wages grow differently for different workers

- Little evidence that quantifies empirically this contribution

Research questions

1. How much do coworkers contribute to future wages?
2. What are the channels through which this contribution is identified?

Literature review

Coworker quality and wage levels

- Specific workplace (via field experiments) [e.g., Mas and Moretti (2009)]
- Local labor market (using population data) [e.g., Cornelissen et al. (2017); Battisti (2017)]

Coworker quality and wage growth

- Herkenhoff et al. (2018); Jarosch et al. (2019); Nix (2020)
- Limitation: use observables (wage or education) as a measure of quality

Firms as learning environments

- Gregory (2019); Arellano-Bover and Saltiel (2021)

Estimation

- Arcidiacono et al. (2012)

Contribution

- Go beyond contemporaneous effects and analyze dynamics

Contribution

- Go beyond contemporaneous effects and analyze dynamics
- Provide evidence on the channels that identify this contribution
  1. Peers joining the firm
  2. Peers leaving the firm
  3. Worker moving into different peer group

Effect of peer quality on future wages

Data: Veneto Worker History

Worker records

- $\sim$ 3 million private-sector workers from 1982 to 2001 $\rightarrow$ entire working history
- information on gender, birth year, region or country of birth

Data: Veneto Worker History

Worker records

- $\sim$ 3 million private-sector workers from 1982 to 2001 $\rightarrow$ entire working history
- information on gender, birth year, region or country of birth

Firm records

- all firms where any worker has worked across Italy
- information on tax number, location, start and closure dates, industry

Data: Veneto Worker History

Worker records

- $\sim$ 3 million private-sector workers from 1982 to 2001 $\rightarrow$ entire working history
- information on gender, birth year, region or country of birth

Firm records

- all firms where any worker has worked across Italy
- information on tax number, location, start and closure dates, industry

Contribution (wage) records

- wages from each labor contract, without top coding
- weeks worked, basic contract info (e.g., full/part-time), qualification (e.g. blue, white)

Summary statistics

Empirical strategy

Baseline regression:

$$w_{i,t+h} = \alpha_i + \beta\bar{\alpha}_{-i,t} + \mathbf{x}'_{it}\gamma + \psi_{jt} + \eta_{ot} + \theta_{oj} + \varepsilon_{it}, \tag{1}$$

- $w_{i,t+h}$ is the log weekly earnings at time $t + h$, where $h \geq 0$

Empirical strategy

Baseline regression:

$$w_{i,t+h} = \alpha_i + \beta \bar{\alpha}_{-i,t} + \mathbf{x}'_{it}\gamma + \psi_{jt} + \eta_{ot} + \theta_{oj} + \varepsilon_{it}, \tag{1}$$

- $w_{i,t+h}$ is the log weekly earnings at time $t + h$, where $h \geq 0$
- α_i is the worker fixed effect
- $\bar{\alpha}_{-i,t}$ is the *average* coworker's quality at time t

Empirical strategy

Baseline regression:

$$w_{i,t+h} = \alpha_i + \beta\bar{\alpha}_{-i,t} + \mathbf{x}'_{it}\gamma + \psi_{jt} + \eta_{ot} + \theta_{oj} + \varepsilon_{it}, \tag{1}$$

- $w_{i,t+h}$ is the log weekly earnings at time $t+h$, where $h \geq 0$
- α_i is the worker fixed effect
- $\bar{\alpha}_{-i,t}$ is the *average* coworker's quality at time t
  - Peer group: workers employed in same firm & occupation in a year.

Empirical strategy

Baseline regression:

$$w_{i,t+h} = \alpha_i + \beta \bar{\alpha}_{-i,t} + \mathbf{x}'_{it}\gamma + \psi_{jt} + \eta_{ot} + \theta_{oj} + \varepsilon_{it}, \tag{1}$$

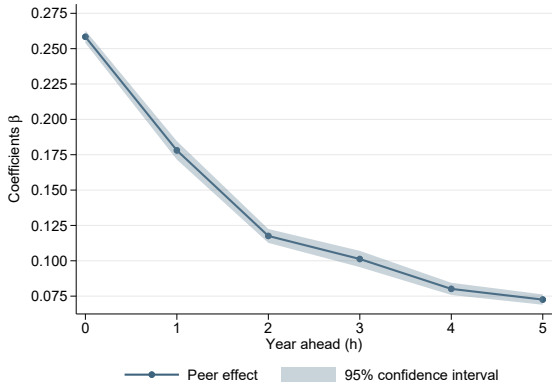
- $w_{i,t+h}$ is the log weekly earnings at time $t + h$, where $h \geq 0$
- α_i is the worker fixed effect
- $\bar{\alpha}_{-i,t}$ is the *average* coworker's quality at time t
  - Peer group: workers employed in same firm & occupation in a year.
- $\mathbf{x}_{it}$ is a set of individual time-varying characteristics
- ψ_{jt}, η_{ot}, θ_{oj} are firm-year, occupation-year, firm-occupation fixed effects

Details on estimation | Summary stats | Distribution change peer effect

Baseline results

Effects of coworker's quality on future wages (β)

$$w_{i,t+h} = \alpha_i + \beta \bar{\alpha}_{-i,t} + \mathbf{x}'_{it}\gamma + \phi_{jt} + \delta_{ot} + \theta_{oj} + \varepsilon_{it},$$



Alternative specification | Heterogeneity 1 | Heterogeneity 2

Mechanisms that identify β: An event-study approach

Discussion on potential mechanisms

The effect of coworker quality β is identified through three mechanisms

1. job stayers: changes in peer when a worker *enters*

Discussion on potential mechanisms

The effect of coworker quality β is identified through three mechanisms

1. job stayers: changes in peer when a worker *enters*
2. job stayers: changes in peer when a worker *leaves*

Discussion on potential mechanisms

The effect of coworker quality β is identified through three mechanisms

1. job stayers: changes in peer when a worker *enters*
2. job stayers: changes in peer when a worker *leaves*
3. job switchers: changes in peer quality when moving to another firm

Discussion on potential mechanisms

The effect of coworker quality β is identified through three mechanisms

1. job stayers: changes in peer when a worker *enters*
2. job stayers: changes in peer when a worker *leaves*
3. job switchers: changes in peer quality when moving to another firm

We provide an event-study analysis of these job changes

Empirical strategy

When a high/low-quality worker enters/leaves:

- Evolution of coworkers' wages in the origin and destination firm when worker joins/leaves
- Compare firms hiring/separating from high- and low-quality workers w/ average-quality workers Diagram
- Ex-ante propensity score matching at the firm-level PS matching hire PS matching leave

Empirical strategy

When a high/low-quality worker enters/leaves:

- Evolution of coworkers' wages in the origin and destination firm when worker joins/leaves
- Compare firms hiring/separating from high- and low-quality workers w/ average-quality workers Diagram
- Ex-ante propensity score matching at the firm-level PS matching hire PS matching leave

When a worker joins high-/low-quality peers:

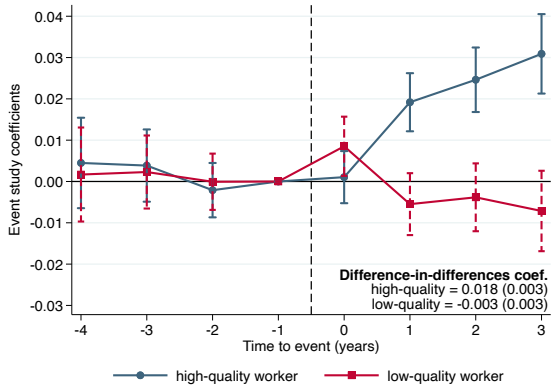
- Evolution of mover's wages when she moves into peer groups of different quality
- Compare workers moving into high- and low-quality peers w.r.t. average-quality peers Diagram
- Ex-ante propensity score matching at the worker-level PS matching

Sample sizes

Event-study results

When a high-/low-quality worker enters

$$w^{new}_{-i,t} = \delta_t + \phi_{j(i)} + \sum_{k \neq -1} \beta_k (Treat_{j(i)} \times \mathbf{1}\{t = k\}) + \epsilon_{-i,t}$$

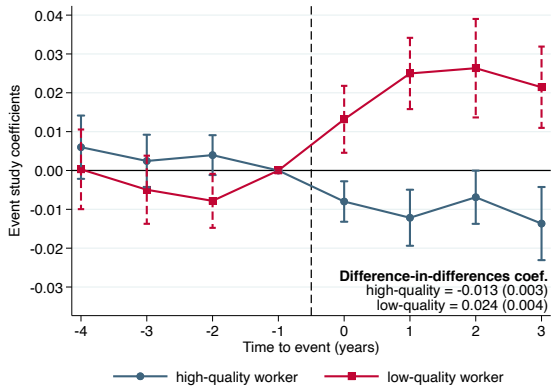


Peers vs non-peers | High- vs low-wage workers | Firms' performance

Event-study results

When a high-/low-quality worker leaves

$$w^{old}_{-i,t} = \delta_t + \phi_{j(i)} + \sum_{k \neq -1} \beta_k (Treat_{j(i)} \times \mathbf{1}\{t = k\}) + \epsilon_{-i,t}$$

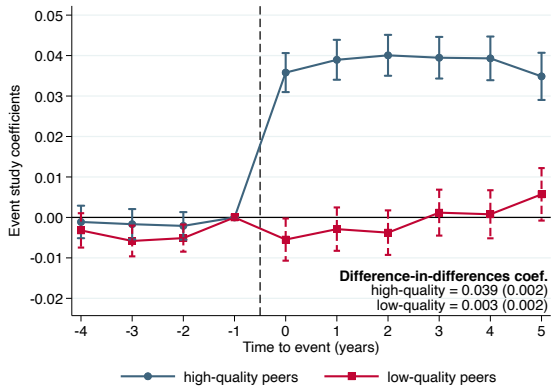


Peers vs non-peers | High- vs low-wage workers

Event-study results

When a worker moves into high-/low-quality peer group

$$w_{i,t} = \delta_t + \eta_i + \sum_{k \neq -1} \gamma_k (Treat_i \times \mathbf{1}\{t = k\}) + \epsilon_{i,t}$$



Blue vs white-collar | High- vs low-wage workers | Continuous treatment

Conclusion

Key takeaways

- Explore an under-studied component of wage growth: coworker quality
- Coworkers play an important role in generating future wages
- Better peers are associated w/higher wages even after 5 years
- Hiring high-quality workers, separating from low-quality workers and moving into high-quality peers imply the highest wage gains

Next steps: Disentangle the mechanisms (structural)

Data: Summary statistics

	(1) Mean	(2) S.D.	(3) Median
Annual earnings	33350.06	40250.33	31730
Weekly wage	744.38	1652.81	652
Weeks worked	42.41	15.24	52
Age	34.54	10.69	32
Tenure	2.45	2.58	2
Firm size	17	75	6
Movers per firm	4	26	1
Peer group size	12	54	4
Mover	0.61	0.49	
Woman	0.36	0.48	
Blue-collar	0.70	0.46	
Manufacturing	0.53	0.50	
Person-year observations		17,723,260	
Number of workers		2,531,411	
Number of firms		168,613	

back

Estimation

We follow the methodology developed by Hong and Sølvsten (in progress)

Rewrite Equation 1 in a matrix form:

$$w = X\delta + C\delta\beta + \varepsilon$$

where

- $w \in \mathbb{R}^n$ and $X, C \in \mathbb{R}^{n \times k}$ are observed: the construction of the coworker matrix C
- $\delta \in \mathbb{R}^k$ and $\beta \in \mathcal{B}$, where $\mathcal{B}$ is compact.

Assumptions

- exogeneity, $\mathbb{E}[\varepsilon | X, C] = \mathbf{0}$,
- homoskedasticity, $\mathbb{E}[\varepsilon\varepsilon' | X, C] = \sigma^2 I_n$ where $\sigma^2 > 0$ is unknown.

Estimation

The non-linear least squares estimator β is as follows.

$$\hat{\beta} = \min_{\beta \in B} Q_n(\beta) = \arg\min_{\beta \in B} \left\{ \min_{\delta \in \mathbb{R}^k} \|w - X\delta - C\delta\beta\|^2 / n \right\} \tag{2}$$

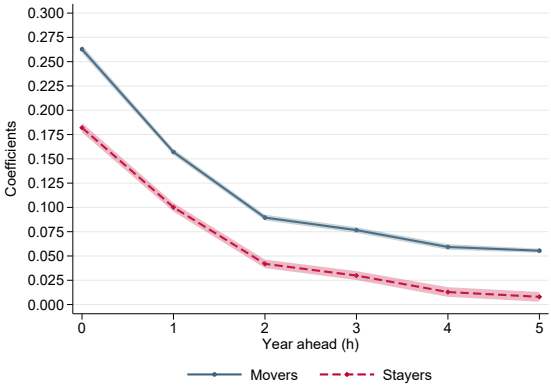
- β is the unique minimizer of sample analogue to $Q_n \to \surd$ consistency (ensured by the exogeneity and homoskedasticity assumptions)
- Taking FOC of δ and β in Equation 2 $\Rightarrow$ moment condition for β

$$S_n(\beta) = w'MC\left(R'R\right)^{-1} R'w/n = 0,$$

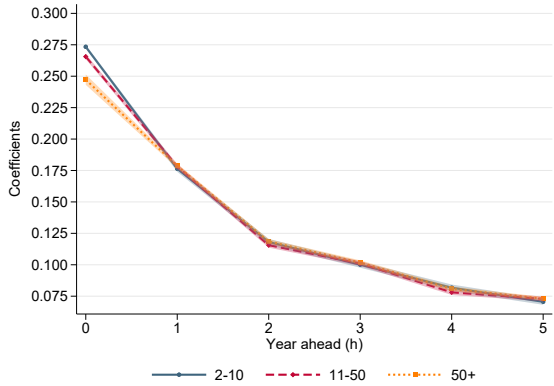
$$R = X + C\beta$$
$$M = I_n - P, \text{ and } P = R\left(R'R\right)^{-1} R'$$

back

Heterogeneous effects/1



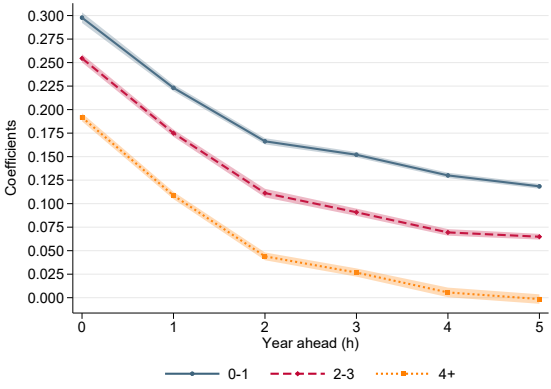
(a) Movers and stayers



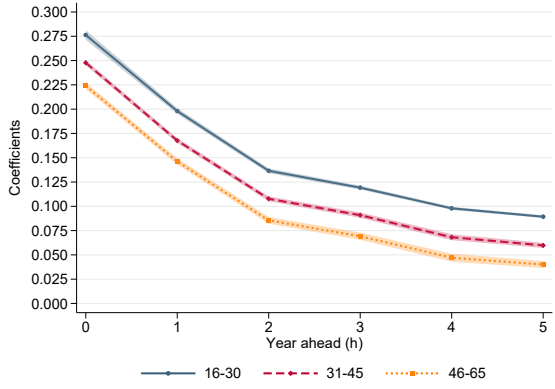
(b) Firm size

back

Heterogeneous effects/2



(a) Tenure brackets

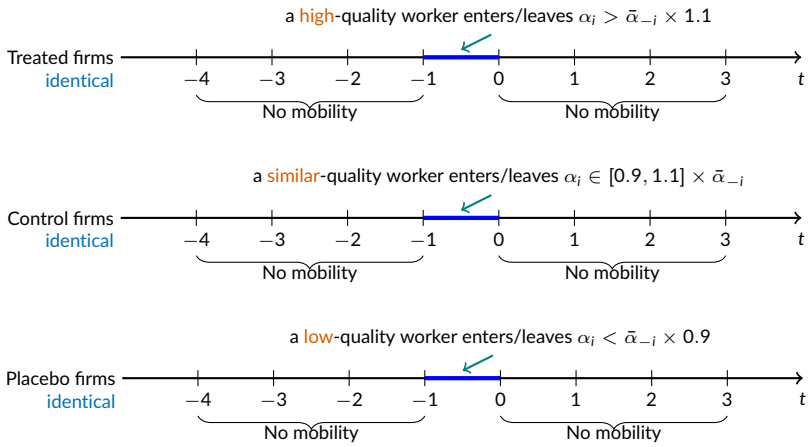


(b) Age brackets

back

Empirical strategy

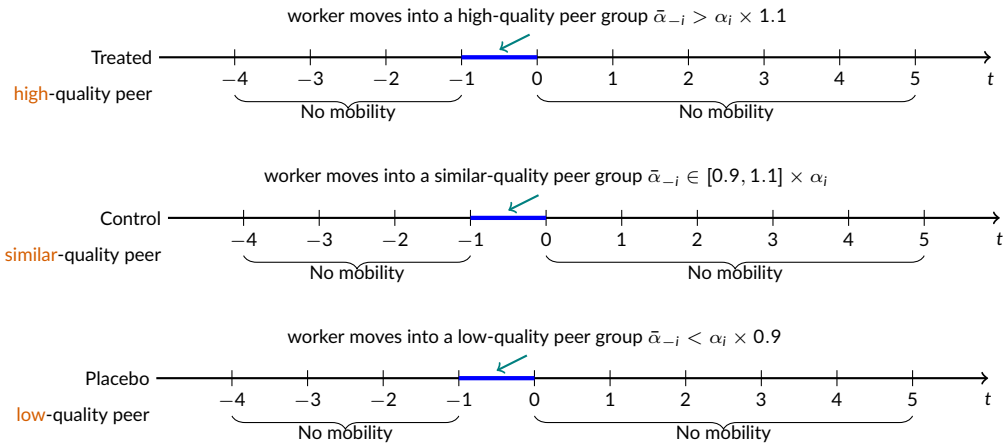
When a high/low-quality worker enters/leaves



back

Empirical strategy

When a worker joins high-/low-quality peers



back

Event-study

Sample sizes

	Treatment	Matched control	Total observations
Panel A: Hiring design	(# of firms)	(# of firms)	(person-year)
High-quality worker	2,164	2,164	285,350
Low-quality worker	1,848	1,848	238,046
Panel B: Leaver design	(# of firms)	(# of firms)	(person-year)
High-quality worker	2,905	2,905	390,135
Low-quality worker	1,885	1,885	234,016
Panel C: Mover design	(# of workers)	(# of workers)	(person-year)
High-quality peers	15,551	15,551	310,220
Low-quality peers	12,778	12,778	255,560

back

The construction of the coworker matrix C

A (simple) example of the coworker matrix C:

- 5 workers, 2 peer groups, and 1 period: data = $\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \\ 4 & 2 \\ 5 & 2 \end{pmatrix}$
- The averaging matrix $\tilde{C} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 0 \end{pmatrix}$
- $C = [\tilde{C}, \mathbf{0}]$ so that C has the same dimension as X.

See the next (backup) slide for more details.

back

The construction of the coworker matrix C

For each worker i at time t (for convenience, we suppress it to i)

$$c_i = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}$$

The averaging matrix $\tilde{C}$ and the coworker matrix C are constructed as follows.

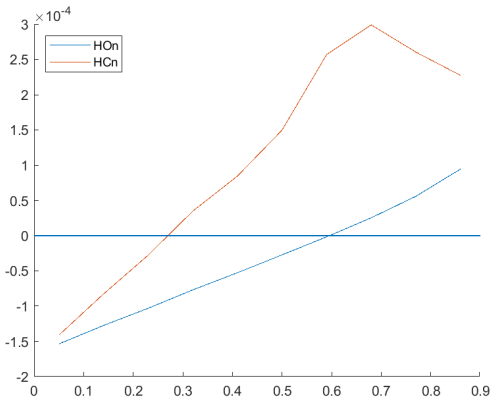
$$\tilde{c}_i = \frac{c_i}{c_i' \cdot \vec{1}} \quad \Rightarrow \quad \tilde{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad \Rightarrow \quad C = \begin{pmatrix} \tilde{C} \\ 0 \end{pmatrix}$$

where the matrix 0 has the same size as F and other covariates

Heteroskedasticity estimator

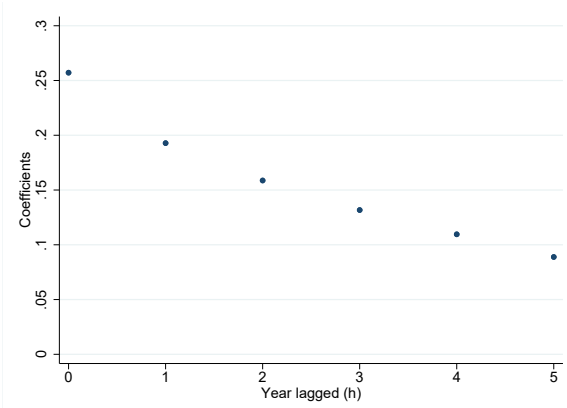
$$w_{it} = \alpha_i + \beta\bar{\alpha}_{-i,t} + \phi_j + \mathbf{x}_{it}'\gamma + \varepsilon_{it},$$

Figure 1: Behaviors of the objective functions of HOn and HCn



Alternative baseline specification

$$w_{i,t} = \alpha_i + \beta \bar{\alpha}_{-i,t-h} + \mathbf{x}'_{it}\gamma + \phi_{jt} + \delta_{ot} + \theta_{oj} + \varepsilon_{it},$$



back

Occupations with low/high learning content

based on Cornelissen et al. (2017) Table F.3

Occupations with low learning content

- Land transport drivers
- Cargo handling and storage
- Laundry, dry cleaning
- Waste disposal
- Packers
- Typists
- Sales persons
- Ready-meal preparers
- Restaurant waiters
- Construction machine attendants
- ...

Occupations with high learning content

- Research in sciences/engineering
- Legal services
- Accounting and tax consultants
- Architects
- Entrepreneurs
- Medical services
- Dentists
- Pharmacist
- Civil engineers
- ...

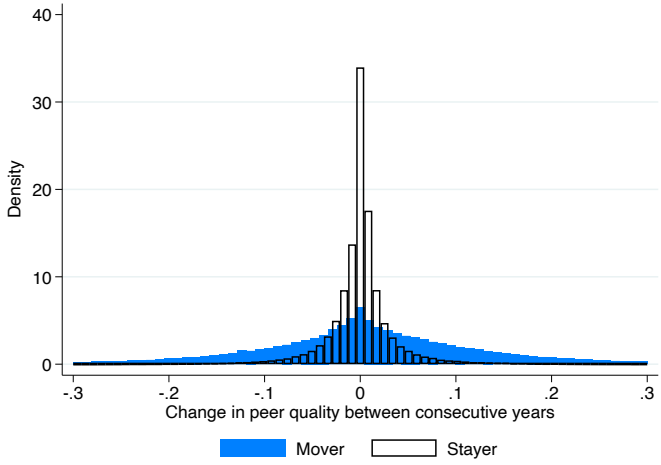
back

Summary statistics after estimation

Statistic	Value
Standard deviation log weekly wages	0.436
Standard deviation worker fixed effect	0.269
Standard deviation peer fixed effect	0.178
Standard deviation occupation-time fixed effect	0.065
Standard deviation firm-occupation fixed effect	0.103
Standard deviation firm-time fixed effect	0.137
Standard deviation change of peer fixed effect between t and $t-1$	0.090
Standard deviation change of peer fixed effect between t and $t-1$ for movers	0.173
Standard deviation change of peer fixed effect between t and $t-1$ for stayers	0.066
Correlation worker fixed effect/peer fixed effect	0.551

back

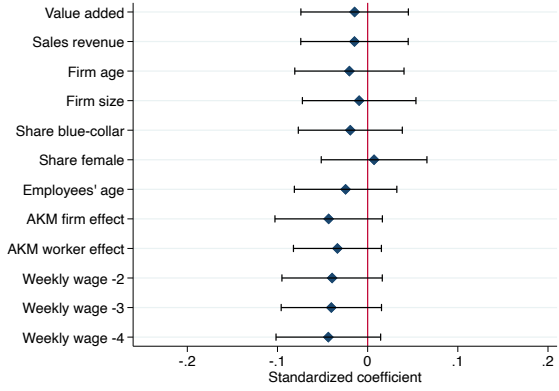
Summary statistics after estimation



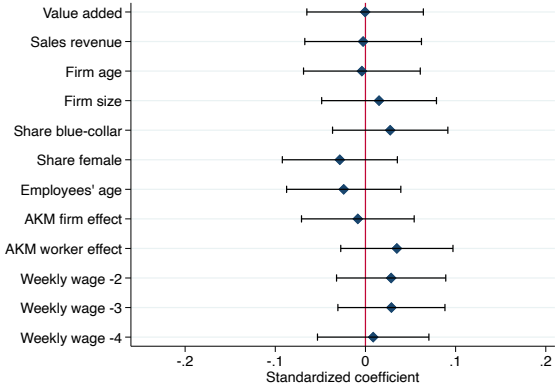
back

Propensity score matching

Balance test of covariates ($t = -3$) - Hiring design



Matched Control and Treatment 1

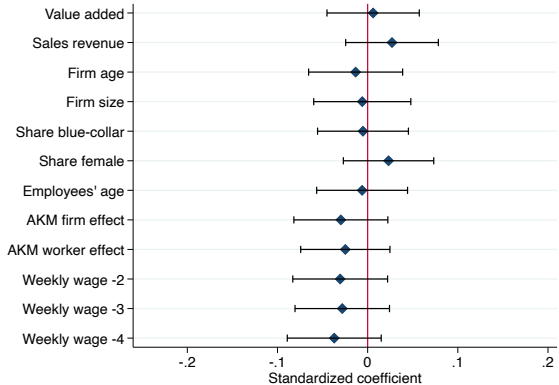


Matched Control and Treatment 2

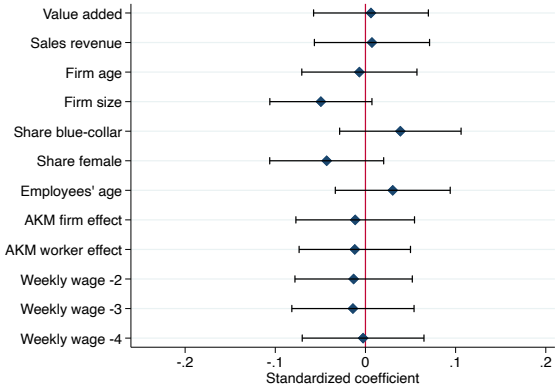
Balance test unmatched | Common support | Back to ES Intro

Propensity score matching

Balance test of covariates ($t = -3$) - Leaver design



Matched Control and Treatment 1

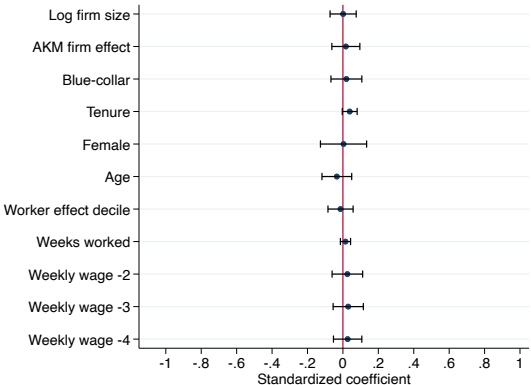


Matched Control and Treatment 2

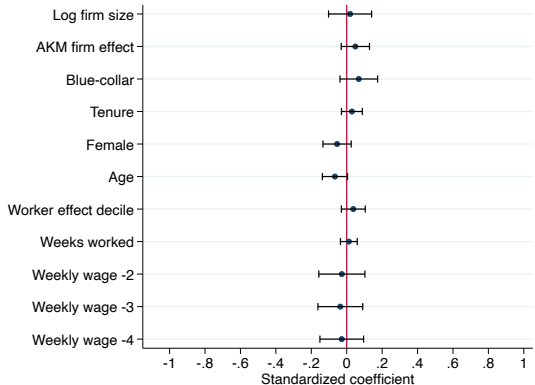
Balance test unmatched | Common support | Back to ES Intro

Propensity score matching

Balance test of covariates ($t = -3$) - Mover design



Matched Control and Treatment 1

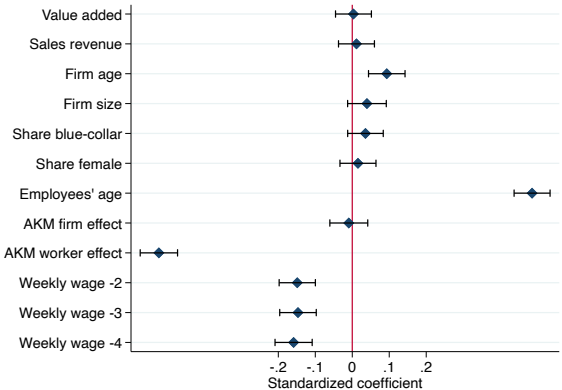


Matched Control and Treatment 2

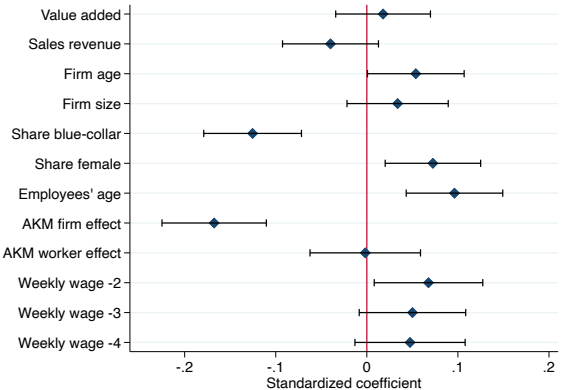
Balance test unmatched | Common support | Back to ES worker

Propensity score matching

Hiring design



[Unmatched Control and Treatment 1]

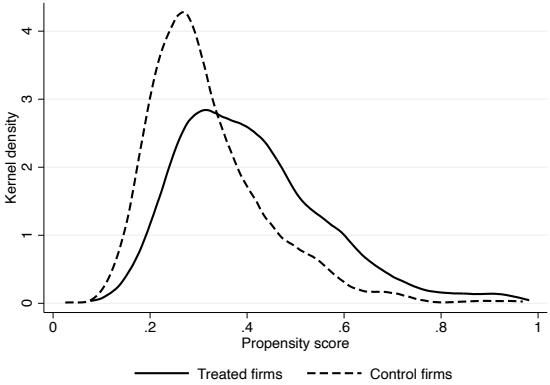


[Unmatched Control and Treatment 2]

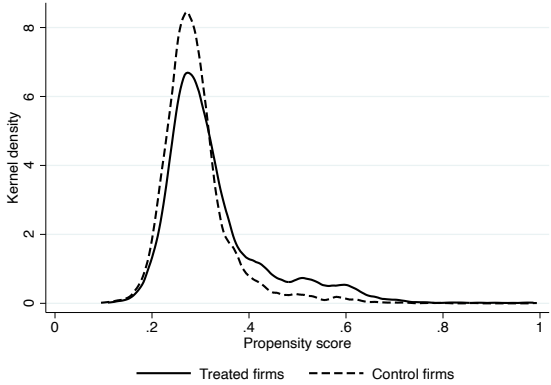
back

Propensity score matching

Hiring design



[Matched Control and Treatment 1]

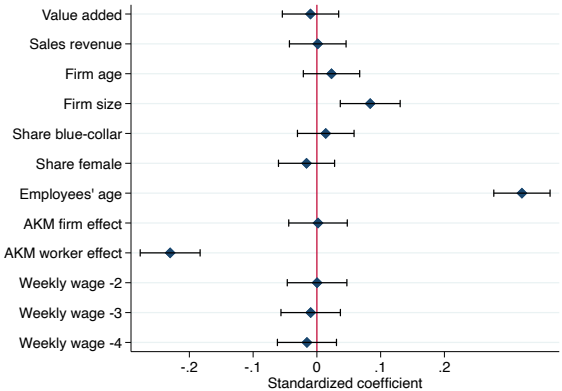


[Matched Control and Treatment 2]

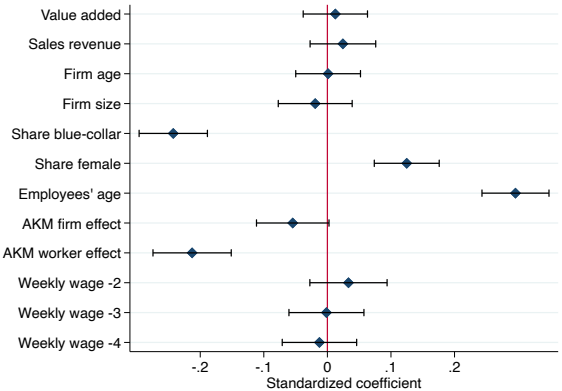
back

Propensity score matching

Leaver design



[Unmatched Control and Treatment 1]

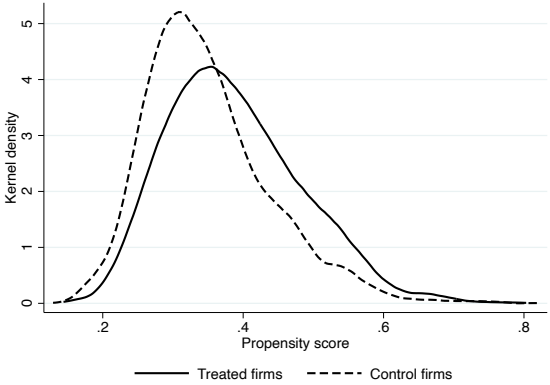


[Unmatched Control and Treatment 2]

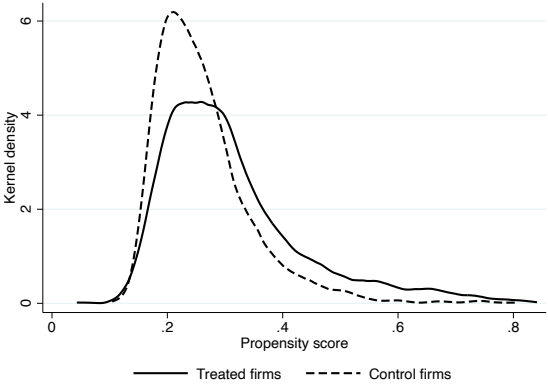
back

Propensity score matching

Leaver design



[Matched Control and Treatment 1]

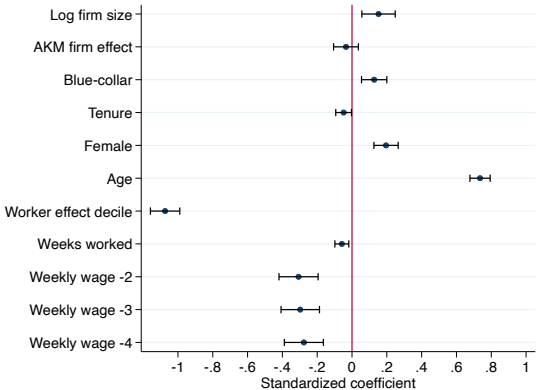


[Matched Control and Treatment 2]

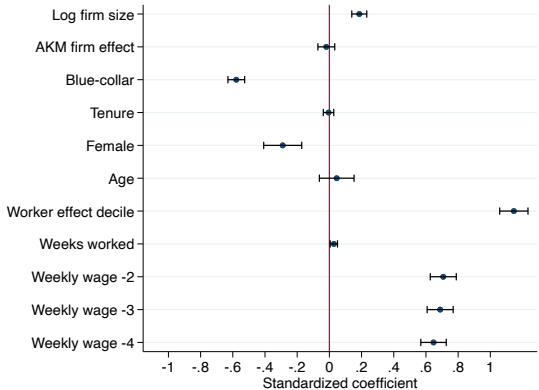
back

Propensity score matching

Worker-level design



[Unmatched Control and Treatment 1]

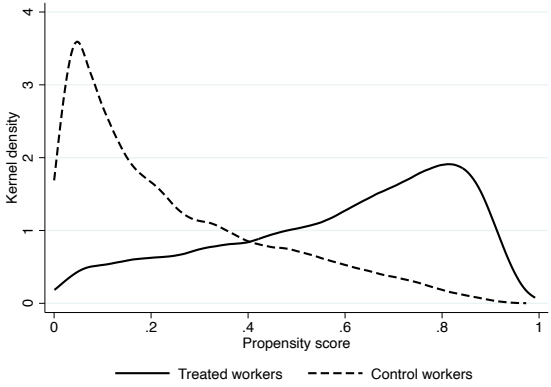


[Unmatched Control and Treatment 2]

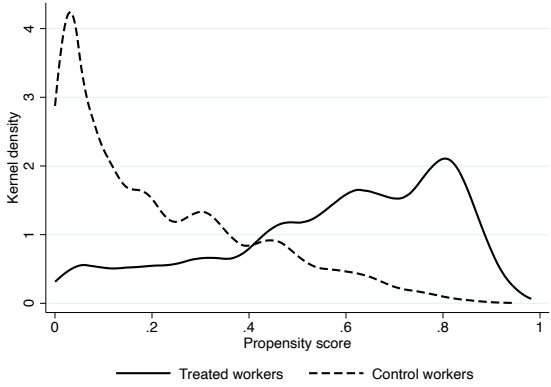
back

Propensity score matching

Worker-level design



[Matched Control and Treatment 1]

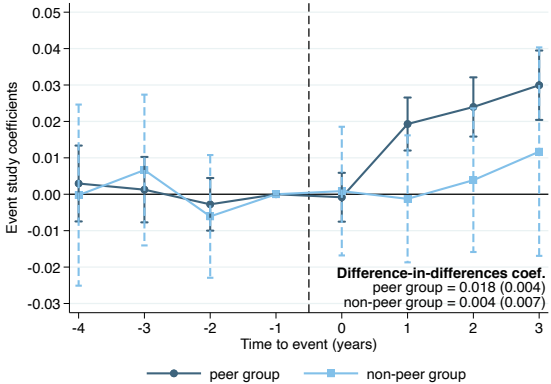


[Matched Control and Treatment 2]

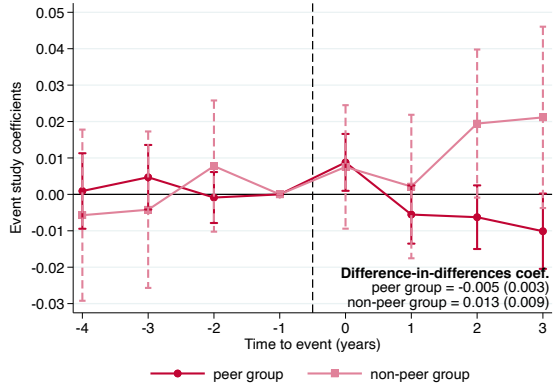
back

Heterogeneity - Hire design

Peers vs non-peers



(a) High-quality entry

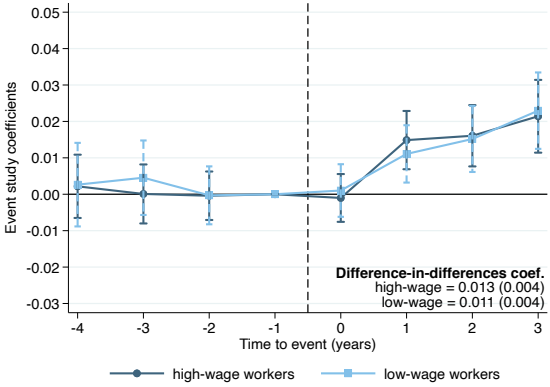


(b) Low-quality entry

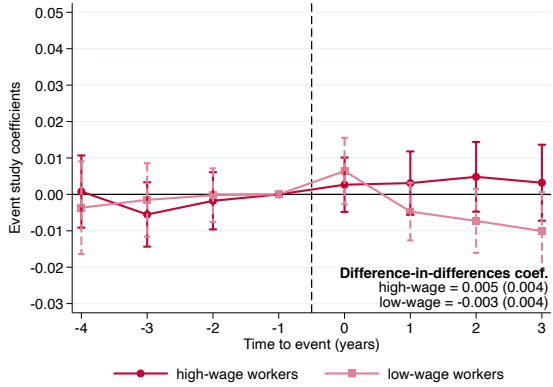
back

Heterogeneity - Hire design

High- vs low-wage workers



(c) High-quality entry

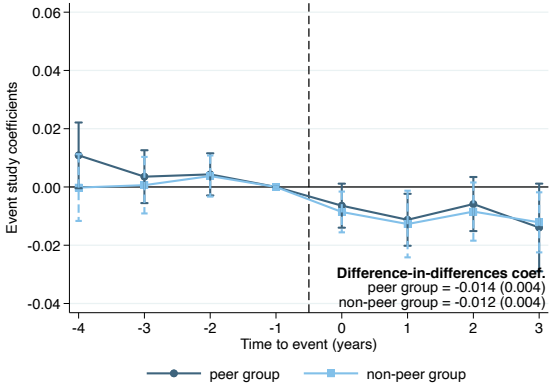


(d) Low-quality entry

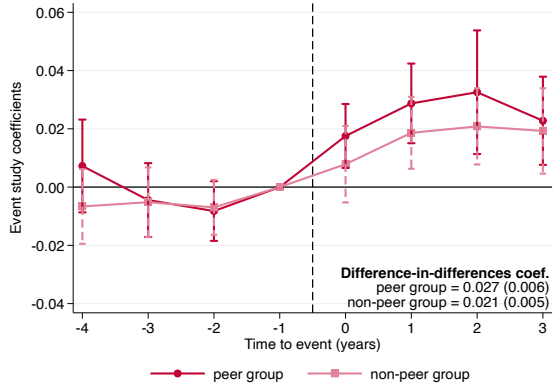
back

Heterogeneity - Leaver design

Peers vs non-peers



(a) High-quality leave

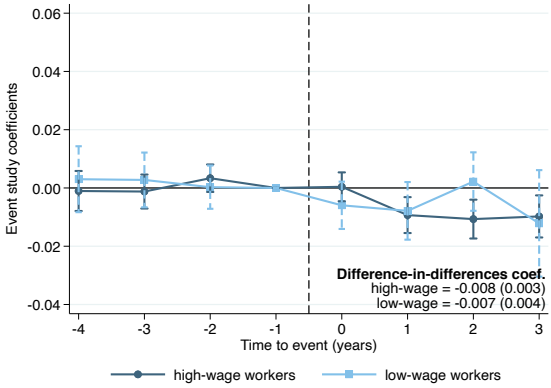


(b) Low-quality leave

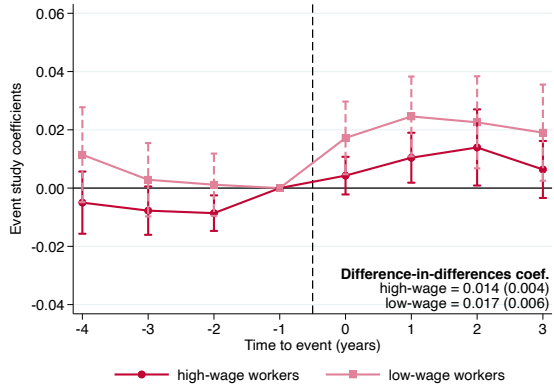
back

Heterogeneity - Leaver design

High- vs low-wage workers



(c) High-quality leave

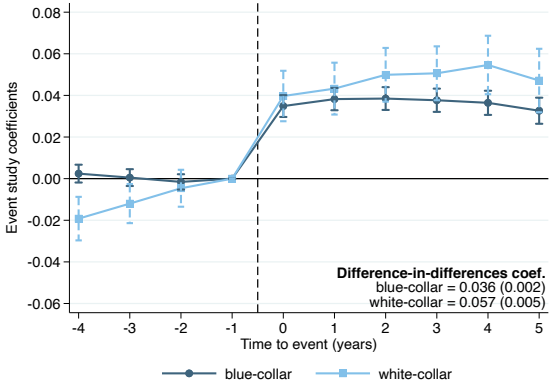


(d) Low-quality leave

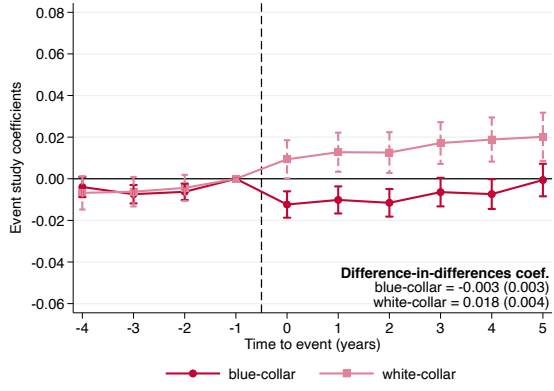
back

Heterogeneity - Mover design

Blue- vs white-collar workers



(a) High-quality mover

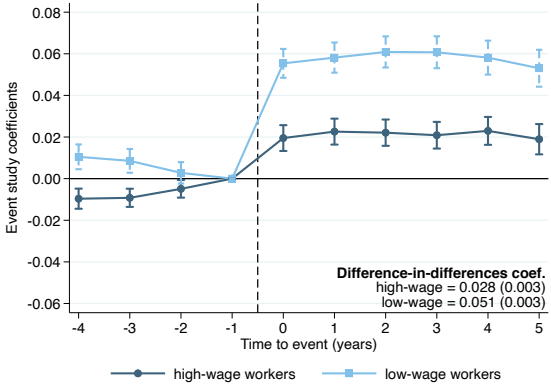


(b) Low-quality mover

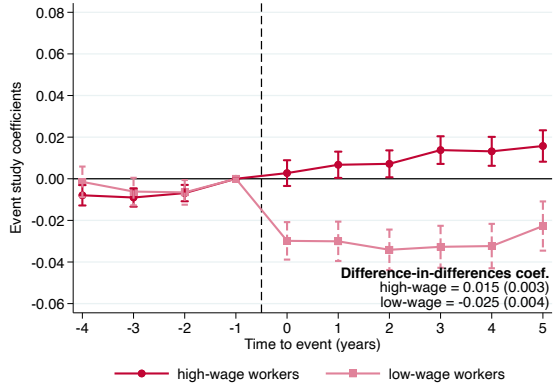
back

Heterogeneity - Mover design

High- vs low-wage workers



(a) High-quality mover

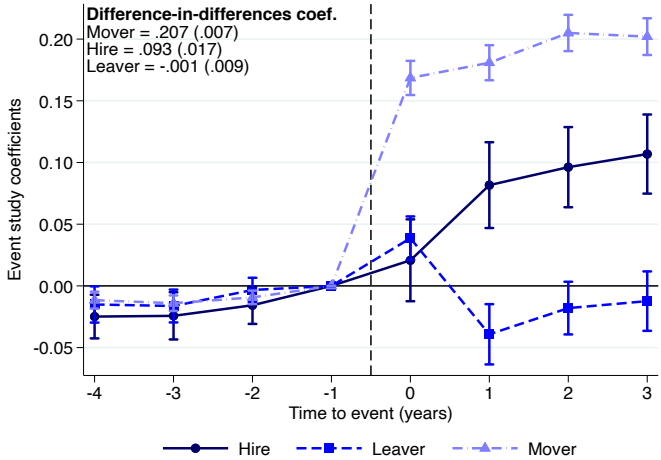


(b) Low-quality mover

back

Event study - Robustness

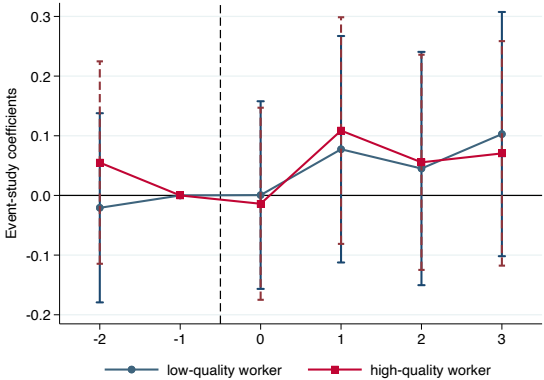
Continuous treatment



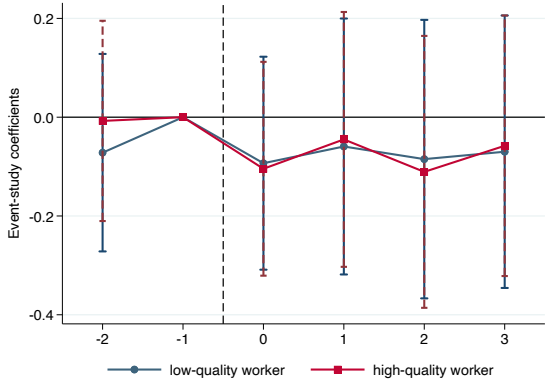
back

Event study - Robustness

Firms' value added and sales



(a) Sales per worker



(b) Value added per worker

back