Online Appendix

Butschek, Sebastian, and Jan Sauermann. 2022. "The Effect of Employment Protection on Firms' Worker Selection"

A AKM person effects

A.1 Estimating AKM person effects

To estimate time-invariant individual productivity we use individual-level spell data on employment and monthly wages (unadjusted for working time) from JOBB.[1] To have sufficient AKM worker effect coverage of new hires for both the period before the introduction of the reform and after, we estimate the AKM model in rolling time periods, i.e. for individuals hired in t, we estimate their AKM worker effect using the employment history for the preceding 7 years (from $t-7$ to $t-1$). In our AKM estimation we include all employment information for individuals aged 18-65 for the years starting in 1986 (the beginning of records) through 2004.

We deflate wages using the CPI and winsorize at 0.5% and 99.5% of the annual real monthly wage distribution. We drop singletons, i.e., individuals that are only observed once in the respective 7-year estimation window. We residualize wages by age, gender and education using the following specification:

$$\begin{aligned}\ln(w_{it}) = \alpha_i &+ \beta_1 male_i + \beta_2 age_{it} + \beta_3 educ_{it} \\ &+ \beta_4 male_i * age_{it} + \beta_5 male_i * educ_{it} \\ &+ \beta_6 age_{it} * education_{it} + \beta_7 male_i * age_{it} * educ_{it} + \epsilon,\end{aligned} \tag{3}$$

where gender is a dummy, age is in years and education is a categorical variable with 8 categories, one of which is missing education information.[2] Information on education is available only from 1990; for earlier years, we impute it from the first year it is available.[3]

[1] Unlike the Wage Statistics Survey (WSS), the JOBB register does not contain information on working hours (or at least full-time status). Using JOBB rather than WSS to estimate an AKM model comes at the cost of lowering the predictive power of AKM worker effects for, e.g., cognitive test scores (Butschek and Sauermann, 2019). However, as WSS is available only for the employees of a changing sub-set of firms, using JOBB gives us an AKM worker effect estimate for a much larger share of new hires in our sample.

[2] 1: *Folkskola*; 2: *Grundskola*; 3: cat. 2 plus 1 or 2 years of *Gymnasieskola*; 4: cat. 2 plus 3 years of *Gymnasieskola*; 5: cat. 3 or cat. 4 plus less than 3 years of *Eftergymnasial utbildning*; 6: cat. 4 plus 3 or more years of *Eftergymnasial utbildning*; 7: cat. 6 plus PhD; 8: information on education is missing.

[3] We do not have reliable information on whether individuals are primarily studying. We therefore omit this information.

Building on Abowd, Kramarz and Margolis (1999) and Card, Heining and Kline (2013) we estimate a two-way fixed-effects regression:

$$\ln(w_{ijt}) = \alpha_i + \psi_j + \gamma_t + x'_{it}\beta + r_{ijt}, \tag{4}$$

where $\ln(w_{ijt})$ is the residualized natural logarithm of individual i's monthly wage at firm j in year t.[4] Moreover, there are additive fixed effects for individuals (α_i) and firms (ψ_j) as well as a set of year dummies (γ_t) and a vector of time-varying individual-level controls (x_{it}). We deviate from Card, Heining and Kline (2013) by including age-group specific year effects as individual-level controls (rather than age squared and age cubed as well as education categories interacted with the year dummies, age squared and age cubed). We use 10 age groups; this allows us to flexibly control for the differential effect of macro shocks on different age groups, while the residualization should take care of life-cycle and gender differences. We also proceed differently than Card, Heining and Kline (2013) by estimating the two-way fixed-effects regression for men and women together so $\hat{\alpha}_i$ is comparable across gender.[5]

One limitation of the AKM approach is that it ignores match quality between worker and firm. It has been shown that this match component makes up around 25% of AKM worker effects (Jackson, 2013; Woodcock, 2015). In this paper we rely on the following assumption: if new hires' AKM worker effects contain a match component from previous employment spells, the resulting error with which we measure their ability is not correlated with the treatment status of the firm that is hiring them.

A.2 Estimation results of AKM estimation

Table A1 provides information on the overall number of workers and firms in the population for each estimation window for which the AKM effects are separately estimated. On average we obtain worker fixed-effect estimates $\hat{\alpha}_i$ for 91.6% of the workers for whom we observe a monthly wage and estimates of ψ_j for the 85.0% of firms that have multiple workers and are connected by worker mobility. Our estimation procedure drops all singletons from the population (8.3% of all workers). Columns (7) to (10) show the share of movers by firm fixed effects quartile.

[4] We residualize rather than including age directly as age is perfectly collinear with the combination of individual effects and year dummies.

[5] To estimate Equation (4), we use `reghdfe` (Correia, 2016). Singletons are not included in the estimation.

Table A1: Descriptive statistics AKM estimation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample		Workers			Firms	Share largest	Share of movers by firm FE quartile			
frame	Overall	Singletons	Worker effects	Overall	Firm effects	connected set	Q1	Q2	Q3	Q4
1986–1992	5,435,319	401,754	5,027,152	356,284	299,109	.9947	.6925	.4464	.5908	.5887
1987–1993	5,422,789	396,617	5,020,275	367,475	310,643	.9940	.7036	.4629	.5494	.5936
1988–1994	5,435,506	417,911	5,011,443	380,430	321,403	.9934	.6932	.5167	.5263	.6050
1989–1995	5,447,014	434,153	5,006,559	390,229	331,166	.9926	.6794	.556	.4825	.6153
1990–1996	5,445,220	469,624	4,968,963	395,067	337,528	.9916	.6688	.5738	.452	.5860
1991–1997	5,433,129	503,608	4,922,598	399,915	341,060	.9908	.6563	.5931	.4356	.5741
1992–1998	5,415,569	513,167	4,895,434	397,728	338,377	.9906	.6500	.5828	.4118	.5609
1993–1999	5,419,934	491,487	4,921,917	396,556	338,802	.9906	.6448	.5722	.4478	.5572
1994–2000	5,460,962	473,694	4,980,853	402,424	342,785	.9907	.6492	.5260	.4652	.5695
1995–2001	5,498,288	451,921	5,040,310	403,595	344,756	.9908	.6557	.5223	.4802	.5867
1996–2002	5,528,757	442,109	5,080,866	402,955	343,871	.9908	.6646	.5181	.5092	.5876
1997–2003	5,544,878	425,812	5,113,684	399,729	341,806	.9907	.6631	.5179	.5117	.5904

Note: the table shows descriptive statistics from the estimation of AKM worker and firm effects. The sample frame relates to the 7 years preceding t (from $t-7$ to $t-1$) that are used for the estimation. Columns (1) to (5) show the total number of workers (1) and firms (4) in the data, the number of singleton (workers) that are dropped for the estimation (2), and the number of estimated AKM worker and firm effects ((3) and (5)). Column (6) shows the share of workers within the largest connected set among workers with AKM worker effects. Columns (7) to (10) show the share of movers by firm effect quartile. Largest connected set and share of movers are identified by using `felsdvreg` (Cornelissen, 2008).

B Quantile-regression analysis

As compared to Ordinary Least Squares, quantile regression, introduced by Koenker and Bassett Jr (1978), makes it possible to study the association between a variable of interest and various positions in the conditional distribution (conditional quantiles) of the outcome variable (instead of the conditional mean). Rather than minimizing the sum of squared residuals, quantile regression minimizes the sum of residuals' weighted absolute values. For instance, for quantile regression for P5, the weight is 0.05 when the residual is positive, i.e., when a data point is above the P5 quantile regression line; and the weight is 0.95 when the residual is negative (see, e.g., Koenker and Hallock, 2001; Safer, Suaray and Watson, 2011, for an introduction to "traditional" (conditional) quantile regression).

More recent work has emphasized that, in the presence of covariates other than the regressor of interest ("treatment"), the interpretation of conditional quantile regression ceases to be straightforward (and, often, ceases to be useful) as it tells us about associations between the treatment and certain quantiles of the outcome variable not just conditional on the treatment, but also conditional on all other covariates (see, e.g. Wenz, 2019, for a discussion of interpretation pitfalls). Based on this insight, methods for estimating unconditional quantile regressions have been developed, which allow the researcher to return to a straightforward interpretation of estimates as the changes at a given quantile of the outcome's unconditional distribution (or of the distribution conditional only on the treatment—see, e.g., Firpo, 2007; Firpo, Fortin and Lemieux, 2009; Powell, 2020). We follow the approach of Powell (2020), which distinguishes explicitly between treatments and additional covariates to provide estimates of quantile treatment effects that are conditional only on the treatment; this allows us to estimate a similar DiD specification as our main one in Equation (1).[6]

Our unconditional quantile regression analysis in Section 6.3 estimates the following DiD specification:

$$y_{it} = \alpha + \beta TR_{ijt} * POST_t + \gamma_0 TR_{ijt} + \gamma_1 Post_t + \delta_0 X_{ijt} + \delta_1 TR_{ijt} * X_{ijt} + \epsilon_{ijt}, \tag{5}$$

where y_{it} is some, potentially time-varying individual-level outcome for a newly hired worker i in year t; $TR_{ijt}1$ is one when firm j hiring i in year t has 2 to 10 employees at the time of hire and zero when it has 11 to 19 workers. Observations between 2001 and 2004 make up the post-reform period, $POST_t = 1$. Observations from 1993 through 1999 are defined as the pre-reform period. The year 2000, when the reform was not yet in effect but during part of which anticipation behavior is possible, is dropped from this DiD specification. X_{ijt} are county-level gross regional product and unemployment rate (for the firm's location) and $TR_{ijt} * X_{jt}$ is their interaction with the treatment dummy (addressing the possibility that macro shocks differentially affect very small and small firms). We use Baker (2016)'s Stata implementation `genqreg` of Powell (2020) to separately estimate Equation (5) for quantiles 0.01, 0.05, 0.1, 0.15, and 0.2.

[6] Several other papers study unconditional quantile regression in difference-in-difference settings (Athey and Imbens, 2006; Melly and Santangelo, 2015; Callaway and Li, 2019).

C Tables

Table C1: Data sources

Data source	Period used	Description	Used for measurement of	Variables used/generated
JOBB	1993-2004	Employment spells	Worker flows	Timing and identity of hire
JOBB	1993-2004	Employment spells	Firm size	Monthly head count
JOBB	1993-2004	Employment spells	Other firm characteristics	Industry, age, location, ownership
JOBB	1993-2004	Employment spells	Worker labor market history	Co-workers
LISA	1990-2004	Various information	Worker characteristics	Highest educational qualification
BAKGRUND	n.a.	Birth register	Worker characteristics	Year of birth, gender
KRIGSARKIVET	1970-2004	Military draft information	Worker ability	Cognitive, psychological test scores
ÅRSKURS9	1988-1997	Grade 9 school grades	Worker ability	GPA at age 15

Table C2: EPL effect on minimum hire quality: analysis and AKM estimation periods

	(1)	(2)	(3)	(4)
Analysis period	1993-2004	1994-2004	1995-2004	1996-2004
DiD estimate: Treated*Post=1	-0.0216***	-0.0209***	-0.0195**	-0.0209**
	(0.0076)	(0.0078)	(0.0080)	(0.0085)
Observations	876,832	813,258	742,301	669,632
Firms	279,459	268,333	255,826	242,487
Adjusted R^2	0.0126	0.0123	0.0113	0.0104

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. Dependent variable is minimum hire quality. Hire quality is measured by AKM worker effects estimated in years $t-7$ to $t-1$ for an individual hired in t. Estimates are from a specification with firm FEs and county-level macro controls interacted with the treatment. The year 2000 is excluded to rule out anticipation effects. Each column restricts the estimation sample to the period specified in the header.

Table C3: EPL effect on minimum hire quality with outliers included (non-winsorized)

	(1) OLS	(2) MACRO	(3) FE
DiD estimate: Treated*Post=1	-0.0195***	-0.0268***	-0.0239***
	(0.0055)	(0.0068)	(0.0075)
Observations	915,545	880,254	880,254
Firms	288,873	280,835	280,835
Adjusted R^2	0.0182	0.0188	0.0128

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. Data include outliers, i.e., firms whose number of hires is above the 99.5th percentile. Dependent variable is minimum hire quality. Hire quality is measured by AKM worker effects estimated in years $t-7$ to $t-1$ for an individual hired in t. Estimates are from different versions of the specification in Equation (1): OLS without controls (1), OLS with county-level macro controls interacted with the treatment (2), firm FEs with county-level macro controls interacted with the treatment. The year 2000 is excluded to rule out anticipation effects.

Table C4: EPL effect on different minimum hire quality measures with outliers included (non-winsorized)

	(1) OLS	(2) MACRO	(3) FE	(4) FE
Panel A: GPA at age 15				
Quality measure:	GPA	GPA	GPA	AKM
DiD estimate: Treated*Post=1	-0.0254***	-0.0357***	-0.0687***	-0.0522***
	(0.0070)	(0.0085)	(0.0096)	(0.0104)
Observations	517,486	497,170	497,170	451,757
Firms	217,330	211,215	211,215	199,684
Adjusted R^2	0.0107	0.0158	0.0091	0.0127
Panel B: Cognitive test scores				
Quality measure:	COG	COG	COG	AKM
DiD estimate: Treated*Post=1	0.0062	0.0012	-0.0195**	-0.0261***
	(0.0063)	(0.0077)	(0.0084)	(0.0090)
Observations	639,052	613,744	613,744	567,795
Firms	241,442	234,377	234,377	224,995
Adjusted R^2	0.0085	0.0177	0.0069	0.0129
Panel C: Psychological test scores				
Quality measure:	NONCOG	NONCOG	NONCOG	AKM
DiD estimate: Treated*Post=1	-0.0202***	-0.0286***	-0.0422***	-0.0277***
	(0.0063)	(0.0077)	(0.0091)	(0.0092)
Observations	619,288	594,442	594,442	553,266
Firms	238,094	231,023	231,023	222,200
Adjusted R^2	0.0095	0.0110	0.0063	0.0127

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. Dependent variable is one of three minimum hire quality measures, based on: GPA at age 15 (Panel A), military draft cognitive test scores (Panel B), military draft psychological test scores (Panel C). Estimates are from different versions of the specification in Equation (1): OLS without controls (1), OLS with county-level macro controls interacted with the treatment (2), firm FEs with county-level macro controls interacted with the treatment (3 and 4). As a benchmark, Column (4) reports the estimated effect on minimum hire AKM for the sample of firms used in the respective panel. The year 2000 is excluded to rule out anticipation effects.

Table C5: EPL effect on minimum hire quality: firm size corrections

	(1)	(2)	(3)
DiD estimate: Treated*Post=1	-0.0216***	-0.0219***	-0.0225***
	(0.0076)	(0.0076)	(0.0072)
Observations	876,832	865,181	1,021,154
Firms	279,459	273,161	344,822
Adjusted R^2	0.0126	0.0126	0.0119

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. The columns show different versions of the reform bite measure: (1) head count -1 all firms; (2) head count - managers and family where known, -1 otherwise; (3) unadjusted head count. Dependent variable is minimum hire quality. Hire quality is measured by AKM worker effects estimated in years $t-7$ to $t-1$ for an individual hired in t. Estimates are from a specification with firm FEs and county-level macro controls interacted with the treatment. The year 2000 is excluded to rule out anticipation effects. Each column restricts the estimation sample to the period specified in the header.

Table C6: EPL effect on minimum hire quality and other percentiles

	(1)	(2)	(3)	(4)	(5)
Percentile	P01	P05	P10	P15	P20
Panel A: AKM worker effects					
DiD estimate: Treated*Post=1	-0.0246***	-0.0197**	0.0066	0.0199**	0.0285***
	(0.0086)	(0.0086)	(0.0083)	(0.0080)	(0.0076)
Observations	876,832	876,832	876,832	876,832	876,832
Firms	279,459	279,459	279,459	279,459	279,459
Adjusted R^2	0.0126	0.0123	0.0100	0.0069	0.0047
Panel B: GPA at age 15					
DiD estimate: Treated*Post=1	-0.0670***	-0.0660***	-0.0589***	-0.0433***	-0.0324***
	(0.0096)	(0.0096)	(0.0095)	(0.0094)	(0.0092)
Observations	494,273	494,273	494,273	494,273	494,273
Firms	209,895	209,895	209,895	209,895	209,895
Adjusted R^2	0.0089	0.0089	0.0084	0.0074	0.0064
Panel C: Cognitive test scores					
DiD estimate: Treated*Post=1	-0.0192**	-0.0189**	-0.0119	0.0002	0.0095
	(0.0083)	(0.0083)	(0.0083)	(0.0082)	(0.0080)
Observations	610,649	610,649	610,649	610,649	610,649
Firms	233,034	233,034	233,034	233,034	233,034
Adjusted R^2	0.0067	0.0067	0.0063	0.0053	0.0043
Panel D: Psychological test scores					
DiD estimate: Treated*Post=1	-0.0410***	-0.0408***	-0.0325***	-0.0190**	-0.0089
	(0.0092)	(0.0092)	(0.0091)	(0.0090)	(0.0088)
Observations	591,364	591,364	591,364	591,364	591,364
Firms	229,682	229,682	229,682	229,682	229,682
Adjusted R^2	0.0062	0.0061	0.0057	0.0048	0.0037

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. Dependent variable: minimum hire AKM (Panel A), minimum hire GPA (Panel B), minimum hire cognitive test score (Panel C), and minimum hire psychological test score (Panel D). Estimates are from Equation (1) specifications with firm fixed effects, county-level macro controls interacted with treatment and year dummies. The year 2000 is excluded to rule out anticipation effects.

Table C7: Placebo thresholds

	(1) 12-20 vs 21-29	(2) 22-30 vs 31-39	(3) 32-40 vs 41-49	(4) 42-50 vs 51-59	(5) 52-60 vs 61-69	(6) 62-70 vs 71-79
Panel A: grade point average						
DiD estimate: Treated*Post=1	-0.0237	-0.0576***	-0.0291	0.0402	-0.0306	0.0386
	(0.0149)	(0.0205)	(0.0254)	(0.0306)	(0.0363)	(0.0398)
Observations	171,160	83,458	51,304	34,940	25,152	19,294
Firms	67,931	34,318	21,879	15,452	11,596	9,108
Adjusted R^2	0.0116	0.0175	0.0217	0.0245	0.0232	0.0273
Panel B: cognitive test score						
DiD estimate: Treated*Post=1	-0.0037	-0.0173	0.0037	0.0120	-0.0646*	0.0499
	(0.0135)	(0.0190)	(0.0247)	(0.0299)	(0.0356)	(0.0391)
Observations	197,727	93,002	56,001	37,660	26,827	20,345
Firms	72,216	35,878	22,670	15,928	11,911	9,360
Adjusted R^2	0.0096	0.0137	0.0130	0.0126	0.0154	0.0160
Panel C: psychological test score						
DiD estimate: Treated*Post=1	-0.0082	-0.0132	-0.0108	-0.0248	-0.0452	0.0295
	(0.0145)	(0.0207)	(0.0268)	(0.0325)	(0.0378)	(0.0433)
Observations	194,010	91,849	55,497	37,396	26,683	20,280
Firms	71,721	35,751	22,619	15,883	11,877	9,354
Adjusted R^2	0.0096	0.0139	0.0138	0.0144	0.0152	0.0130

Note: *** p<0.01, ** p<0.05, * p<0.1, with standard errors clustered at the firm level. Dependent variable is one of three minimum hire quality measures, based on: GPA at age 15 (Panel A), military draft cognitive test scores (Panel B), military draft psychological test scores (Panel C). Estimates are from Equation (1) specifications with firm fixed effects, county-level macro controls interacted with treatment and year dummies. The year 2000 is excluded to rule out anticipation effects.

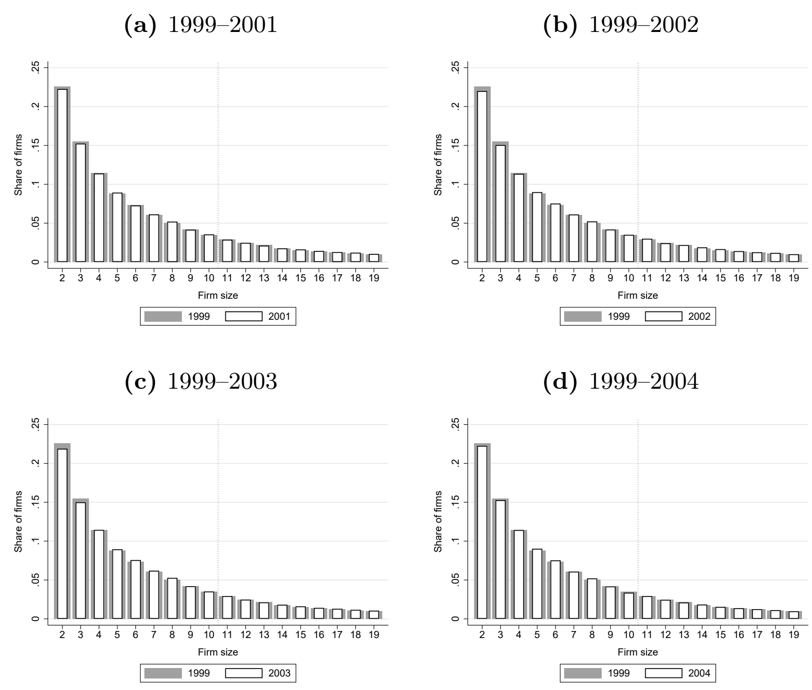
Table C8: Worker-level EPL effect at the bottom of the outcome distribution: quantile regression estimates

Unconditional quantile	(1) P01	(2) P05	(3) P10	(4) P15	(5) P20
Panel A: GPA at age 15					
DiD estimate: Treated*Post=1	-0.0000	-0.0500	-0.0000	-0.0000	0.2500
	(0.0006)	(0.0320)	(0.1267)	(0.3812)	(0.5485)
Observations	1082195	1082195	1082195	1082195	1082195
Panel B: cognitive test scores					
DiD estimate: Treated*Post=1	0.0000	0.0052	0.0026	-0.0024	0.0017
	(0.1576)	(0.3801)	(0.0493)	(0.1344)	(0.3307)
Observations	1459096	1459096	1459096	1459096	1459096
Panel C: psychological test scores					
DiD estimate: Treated*Post=1	-0.0073	-0.0043	-0.0227	-0.0045	-0.0129
	(0.7181)	(0.2783)	(0.1395)	(313.7405)	(0.0801)
Observations	1376541	1376541	1376541	1376541	1376541

Note: *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is estimated GPA at age 15 (Panel A) or, from the military draft, either cognitive test scores (Panel B) or psychological test scores (Panel C). Estimates are from Equation (5), see Appendix Section B for details. The year 2000 is excluded to rule out anticipation effects.

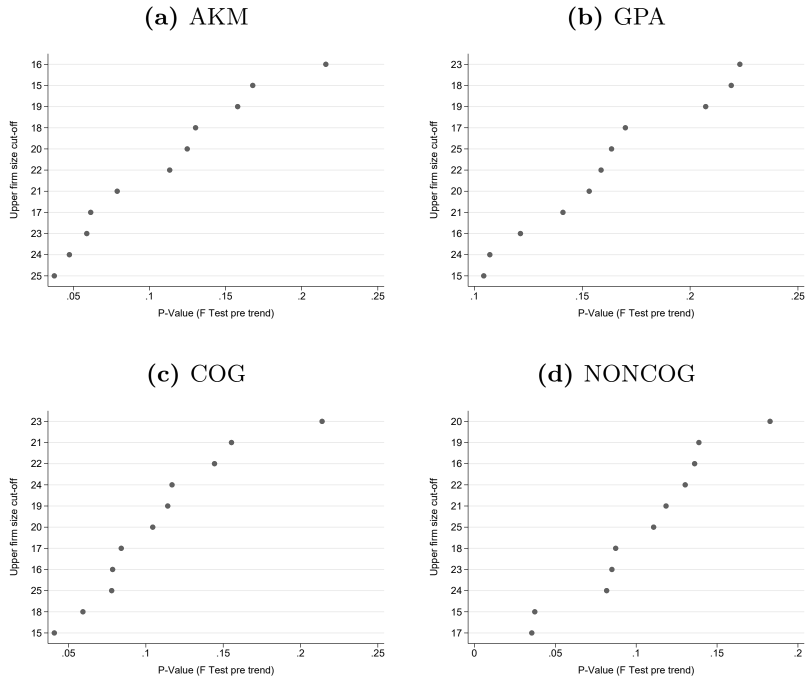
D Figures

Figure D1: Changes in the firm size distribution for different intervals



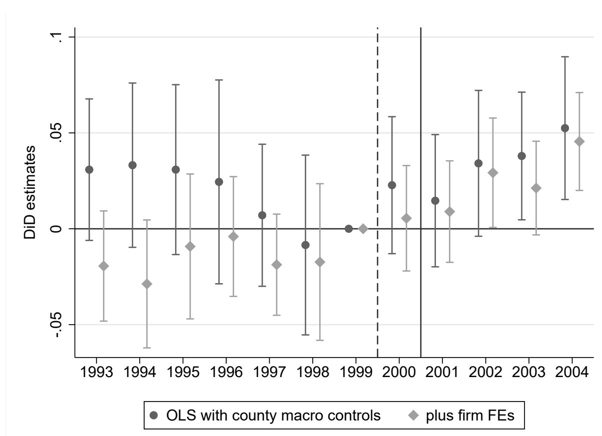
Note: The gray bars show the firm size distribution for firms in 1999. White bars show the corresponding distribution for the years 2001 (a) to 2004 (d).

Figure D2: p values of joint significance tests for potential control groups



Note: For each control group candidate of size 11-15, ..., 11-25, the figures plot the average p value across 50 random draws of an F test of joint significance of $\hat{\beta}_{1993}$, $\hat{\beta}_{1994}$, ..., $\hat{\beta}_{1999}$ from estimating Equation (2).

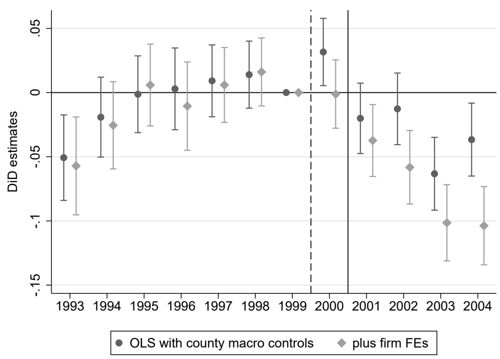
Figure D3: Effects on involuntary separations



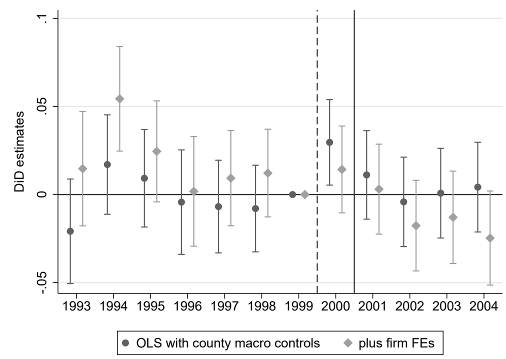
Note: This figure shows yearly DiD estimates for the effect of the EPL reform on the involuntary separation rate (number of involuntary separations divided by previous-year average headcount). Involuntary separations are approximated with separations that are either followed by at least a month of non-employment or by a direct transition to a job that pays less. Estimates are from the specification in Equation (2) with firm fixed effects, county-level macro controls interacted with the treatment and year dummies. Vertical bars denote 95% confidence intervals.

Figure D4: EPL effect on minimum hire quality

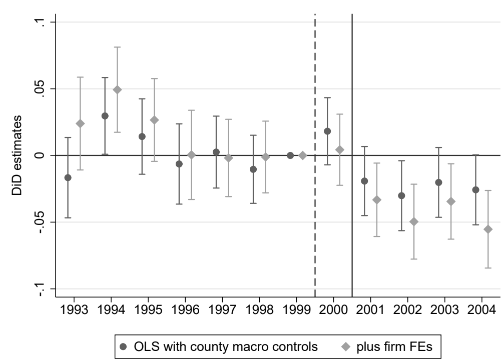
(a) GPA



(b) COG



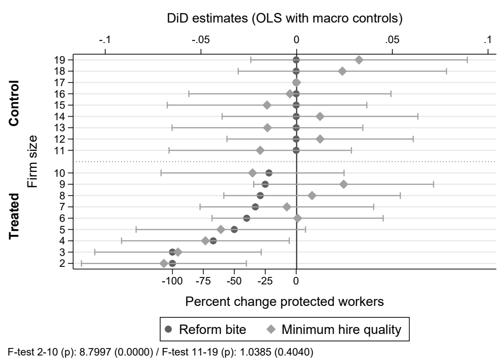
(c) NONCOG



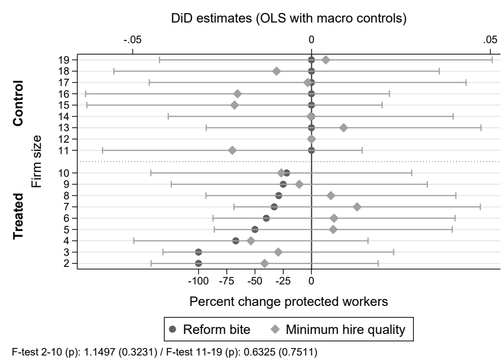
Note: Estimates are from the specification in Equation (2) with firm fixed effects, county-level macro controls interacted with the treatment and year dummies. Hire quality is measured by GPA at age 15 (a) and, from the military draft, cognitive test scores (b) as well as psychological test scores (c). Vertical bars denote 95% confidence intervals. Standard errors are clustered at the firm level.

Figure D5: EPL effect heterogeneity by firm size

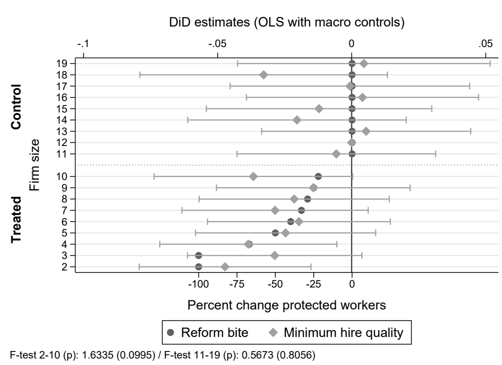
(a) GPA



(b) COG

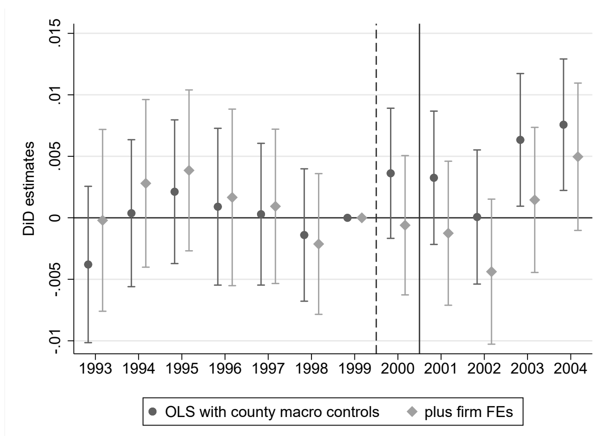


(c) NONCOG



Note: The light gray diamonds show firm size-specific DiD estimates for the effect of the EPL reform on minimum hire quality from an OLS specification (without firm fixed effects) with county-level macro controls interacted with the treatment and year dummies. Hire quality is measured by GPA at age 15 (a) and, from the military draft, cognitive test scores (b) as well as psychological test scores (c). Horizontal bars denote 95% confidence intervals. Standard errors are clustered at the firm level. The dark gray dots illustrate the bite of the reform in terms of the percentage change of workers protected by the seniority rule. The reference firm size category of 12 has been chosen so that it roughly coincides with the mean of estimated firm size-specific DiD coefficients for firms in the control group.

Figure D6: EPL effect on firm attractiveness



Note: This figure shows yearly DiD estimates for the effect of the EPL reform on the share of hires who are voluntary job changers coming directly from high-wage firms (estimated AKM firm effect for 1986-92 above the median). Estimates are from the specification in Equation (2) with firm fixed effects, county-level macro controls interacted with the treatment and year dummies. Vertical bars denote 95% confidence intervals. Vertical bars denote 95% confidence intervals. Standard errors are clustered at the firm level.