APPENDIX - PRODUCTIVITY TRAINING AND WORKER INDEPENDENCE: EVIDENCE FROM INDIA

Niccolò Borri

December 5, 2025

1 Appendix

1.1 Tables

Table A1: Frequency table for Training Type

Training Type	Frequency	Percent
Cost Reduction	38	1.79%
Formal Certifications	654	30.81%
Improve Production Process	1,045	49.22%
Information Technology	56	2.64%
Job Instruction	28	1.32%
Others	11	0.52%
Safety	220	10.36%
Soft Skills	71	3.34%

Note: Data sourced from the Personnel, Production, and Attendance dataset. The training descriptions have been grouped into these 8 macro categories.

Table A2: Regression Results for Rail Production Exact Matching

	Rolling Rate	**Log Blooms per Worker**	**Percent Defected**	**Delay**
Abs_{izt}	-0.085**	-0.258***	0.0001	0.094
	(0.030)	(0.050)	(0.002)	(0.090)
$Post_{izt}$	0.043*	0.120*	0.002	-0.035
	(0.020)	(0.060)	(0.002)	(0.080)
$Abs_{izt} * Post_{izt}$	0.079*	0.247**	-0.003	-0.139
	(0.030)	(0.070)	(0.002)	(0.120)
N	292	292	292	292

Note: Regression results using nearest neighbor matched sample using exact matching on teams, rail shifts between June 2000 and June 2002. * 10%, ** 5% and * 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \eta_i + \theta X_{izt} + \epsilon_{izt}$. γ_z team fixed effect, τ_t month fixed effects, η_i shift fixed effects. First column reports results for output per minute worked. Second column reports results for log output made per worker in a shift. Third column reports results for percentage of defective output. Fourth column reports results for minutes of delay during a shift. Standard errors are clustered at the team by month level and reported in parenthesis.

Table A3: Regression Results for shifts with low planned maintenance

	Rolling Rate	**Log Blooms per Worker**	**Percent Cobbled**	**Delay**
Abs_{izt}	-0.183***	-0.148**	-0.004***	0.071
	(0.05)	(0.05)	(0.00)	(0.07)
$Post_{izt}$	-0.008	0.076*	-0.001	-0.116**
	(0.03)	(0.03)	(0.00)	(0.04)
$Abs_{izt} * Post_{izt}$	0.192***	0.196**	0.001	-0.119
	(0.05)	(0.06)	(0.00)	(0.09)
N	1138	1120	1139	1139

Note: Regression results excluding shifts where more than half of the shifts affected by planned maintenance between June 2000 and June 2002. * 10%, ** 5%, and *** 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \eta_i + \theta X_{izt} + \epsilon_{izt}$. γ_z team fixed effect, τ_t month fixed effects, η_i shift fixed effects. First column reports results for output per minute worked. Second column for log output made per worker in a shift. Third column for percentage of cobbled output. Fourth column for minutes of unexpected delay. Standard errors clustered at the team-by-month level and reported in parenthesis.

1.2 Figures

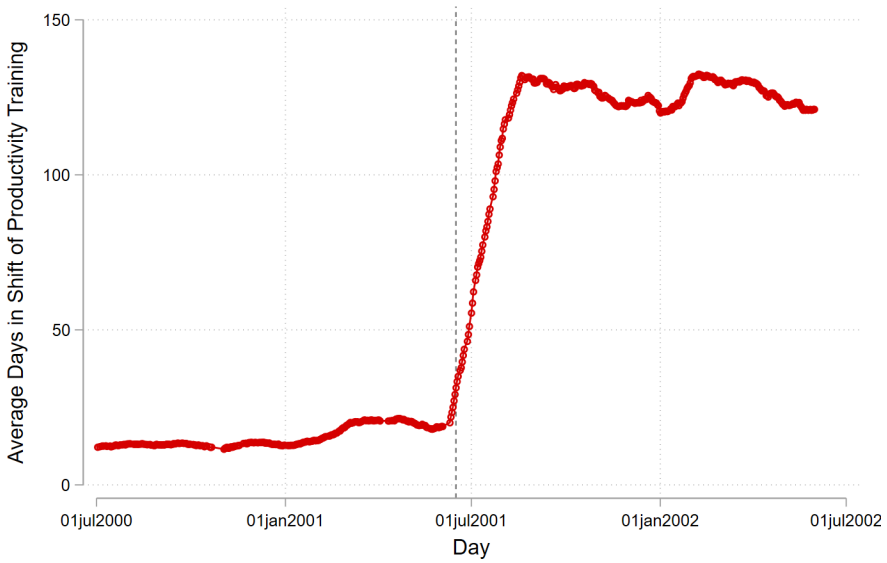

Figure A1: 30 Days Moving Average Productivity Training Days by Shift

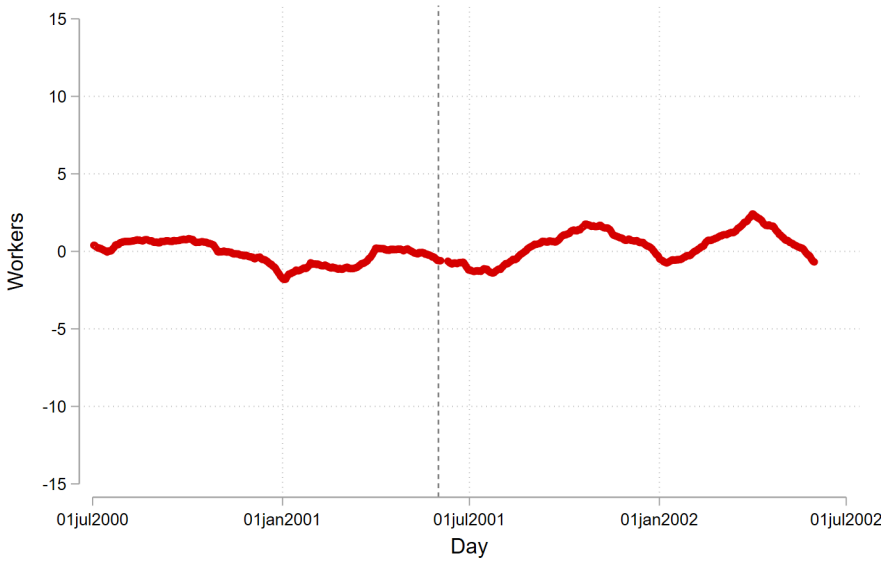

Figure A2: Detrended 90 Days Moving Average Workers in Shifts

1.3 Structural Components

In the appendix, I present results for shifts involved in producing structural components. These results were excluded from the main analysis because the training program did not emphasize practices relevant to this type of production. Additionally, the production of structural components became less significant over time as the focus shifted towards rail output. My estimates indicate that the training had no discernible effect on structural component production, as demonstrated in tables A1 through A4. I estimate equation 1 to assess the impact of training on structural output. figure A1 illustrates the 60-day moving average of structural output considering only shifts making structural components.

Table A4: Regression Results for Structural Production with Weights

	Rolling Rate	**Log Blooms per Worker**	**Percent Defected**	**Delay**
Abs_{izt}	-0.027*	0.067	-0.004	0.190**
	(0.010)	(0.060)	(0.001)	(0.070)
$Post_{izt}$	-0.029*	-0.222**	0.007	-0.269**
	(0.010)	(0.080)	(0.001)	(0.090)
$Abs_{izt} * Post_{izt}$	0.091**	0.059	0.002	0.056
	(0.030)	(0.160)	(0.010)	(0.180)
N	367	367	367	367

Note: Regression results on sample of structural shifts between June 2000 and June 2002, weighted using inverse propensity scores procedure explained in section 5. * 10%, ** 5% and * 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \theta X_{izt} + \epsilon_{izt}$. γ team fixed effect, τ month fixed effects. First column reports results for output per minute worked. Second column reports results for log output made per worker in a shift. Third column reports results for percentage of defective output. Fourth column reports results for minutes of delay during a shift. Standard errors are clustered at the team by month level.

Table A5: Regression Results for Structural Production without Weights

	Rolling Rate	**Log Blooms per Worker**	**Percent Defected**	**Delay**
Abs_{izt}	-0.024	-0.036	0.000	0.068
	(0.020)	(0.070)	(0.000)	(0.090)
$Post_{izt}$	-0.015	-0.120*	0.004	-0.160**
	(0.010)	(0.060)	(0.000)	(0.050)
$Abs_{izt} * Post_{izt}$	0.072	0.202	-0.005	0.115
	(0.050)	(0.240)	(0.010)	(0.190)
N	429	429	429	429

Note: Regression results on sample of structural shifts between June 2000 and June 2002, non weighted. * 10%, ** 5% and * 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \theta X_{izt} + \epsilon_{izt}$. γ team fixed effect, τ month fixed effects. First column reports results for output per minute worked. Second column reports results for log output made per worker in a shift. Third column reports results for percentage of defective output. Fourth column reports results for minutes of delay during a shift. Standard errors are clustered at the team by month level.

Table A6: Regression Results for Structural Production NNM

	Rolling Rate	**Log Blooms per Worker**	**Percent Defected**	**Delay**
Abs_{izt}	-0.016	0.006	-0.005	0.125
	(0.020)	(0.100)	(0.001)	(0.070)
$Post_{izt}$	-0.043	-0.240	0.015	-0.131
	(0.050)	(0.310)	(0.010)	(0.260)
$Abs_{izt} * Post_{izt}$	0.037	0.349	-0.001	-0.036
	(0.050)	(0.300)	(0.010)	(0.280)
N	80	80	80	80

Note: Regression results on nearest neighbor matched sample of structural shifts between June 2000 and June 2002. * 10%, ** 5% and * 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \theta X_{izt} + \epsilon_{izt}$. γ team fixed effect, τ month fixed effects. First column reports results for output per minute worked. Second column reports results for log output made per worker in a shift. Third column reports results for percentage of defective output. Fourth column reports results for minutes of delay during a shift. Standard errors are clustered at the team by month level.

Table A7: Regression Results for Structural Production Exact Matching

	Rolling Rate	**Log Blooms per Worker**	**Percent Defected**	**Delay**
Abs_{izt}	0.003	0.157	-0.004	-0.018
	(0.030)	(0.140)	(0.010)	(0.150)
$Post_{izt}$	-0.016	-0.270	0.007	-0.352*
	(0.060)	(0.300)	(0.010)	(0.140)
$Abs_{izt} * Post_{izt}$	0.032	0.110	-0.010	0.123
	(0.080)	(0.360)	(0.020)	(0.210)
N	80	80	80	80

Note: Regression results on nearest neighbor matched sample using exact matching on teams, structural shifts between June 2000 and June 2002. * 10%, ** 5% and * 1% significance level. Equation: $y_{izt} = \beta_0 + \beta_1 Abs + \beta_2 Post + \beta_3 Post * Ab_{izt} + \gamma_z + \tau_t + \theta X_{izt} + \epsilon_{izt}$. γ team fixed effect, τ month fixed effects. First column reports results for output per minute worked. Second column reports results for log output made per worker in a shift. Third column reports results for percentage of defective output. Fourth column reports results for minutes of delay during a shift. Standard errors are clustered at the team by month level.

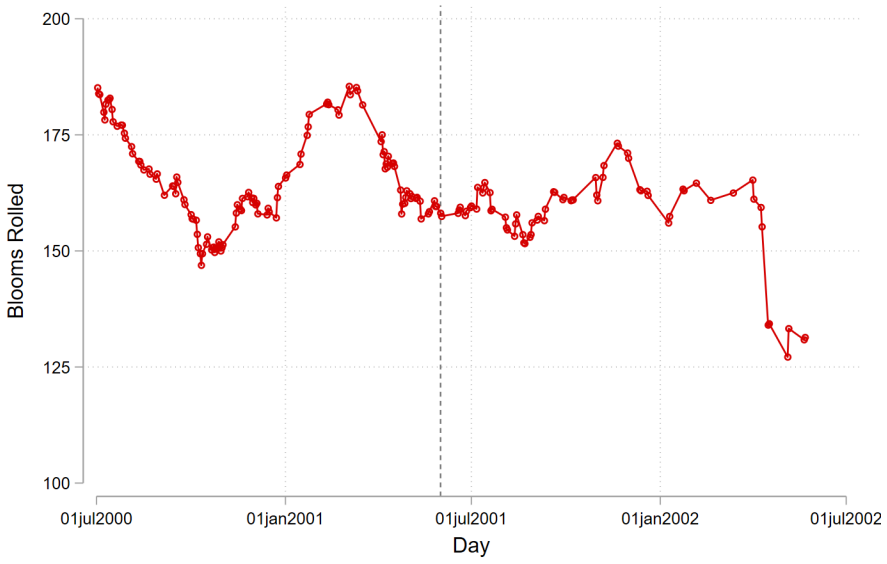

Figure A3: 90 Days Moving Average Structural Production