Machines and Machinists: Incremental Technical Change and Wage Inequality

Miklós Koren   Márton Csillag   János Köllő

Motivation

Motivation

1. Wage inequality in past decades
2. Technological revolutions

Weaver productivity across countries and over time

Clark, 1987

"In 1910 one New England cotton textile operative performed as much work as 1.5 British, 2.3 German, and nearly 6 Greek, Japanese, Indian, or Chinese workers."

Bessen, 2012

"A typical weaver in the United States in 1902 produced over *50 times* as many yards of cloth in an hour of weaving as did a weaver a century earlier producing a comparable cloth."

Not all of it is *quantity* of capital

Bessen, 2012

"The weaver in 1902, however, achieved that output using *eighteen* power-driven looms while the weaver of 1802 used a single handloom."

Sutton, 2001

"On technical performance, there was a small but significant quality gap in favour of the imported [rather than Indian] machine."

Outline

An engineering model of machines and machinists

Standard model

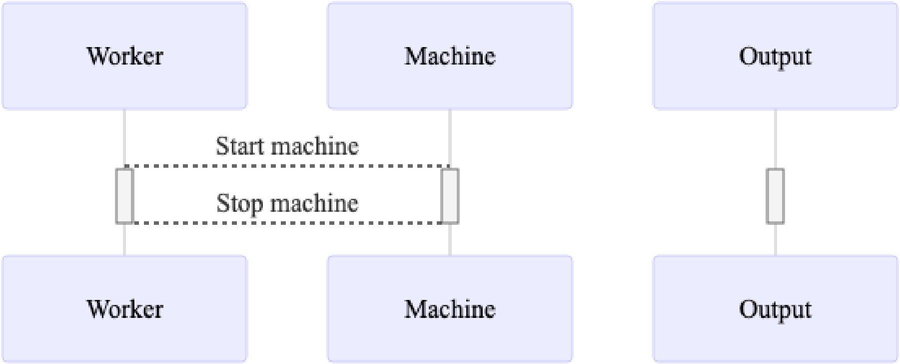
$$Y = AK^{\alpha}L^{1-\alpha}$$

How do machines and people work together?

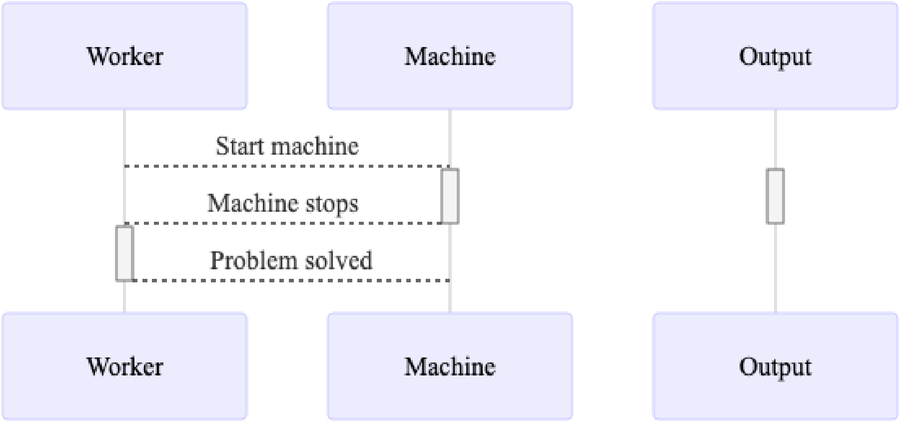
Tool model A worker feeds material into a metal press (both worker and machine busy) to produce.

Operator model A power loom produces in an autonomous fashion (worker idle), until a problem arises. The operator fixes it (machine idle) to get it back to work as fast as possible.

Tool model


Worker
Machine
Output
Start machine
Stop machine
Worker
Machine
Output

Operator model



Two measures of quality

machine quality Expected autonomous uptime θ
worker quality Speed of fixing problems h

Machine busy for θ, idle for $1/h$.

Expected fraction of time working : $\theta h/(1+\theta h) \equiv x$.

Worker busy $(1-x)$ fraction of the time.

Production function

$$dY = \begin{cases} Adt & \text{if machine running, } s = 1 \\ 0 & \text{if not, } s = 0 \end{cases}$$

Markov chain for machine uptime

Kolmogorov equation:

$$\dot{\pi}_1(t) = -\frac{1}{\theta}\pi_1(t) + h\pi_0(t).$$

Ergodic distribution

$$\frac{1}{T}\int_{t=0}^{T}\pi_1(t)dt \approx \pi_1^*.$$

The steady-state probability is the solution to $-\frac{1}{\theta_m}\pi_1(t) + h_i\pi_0(t) = 0$,

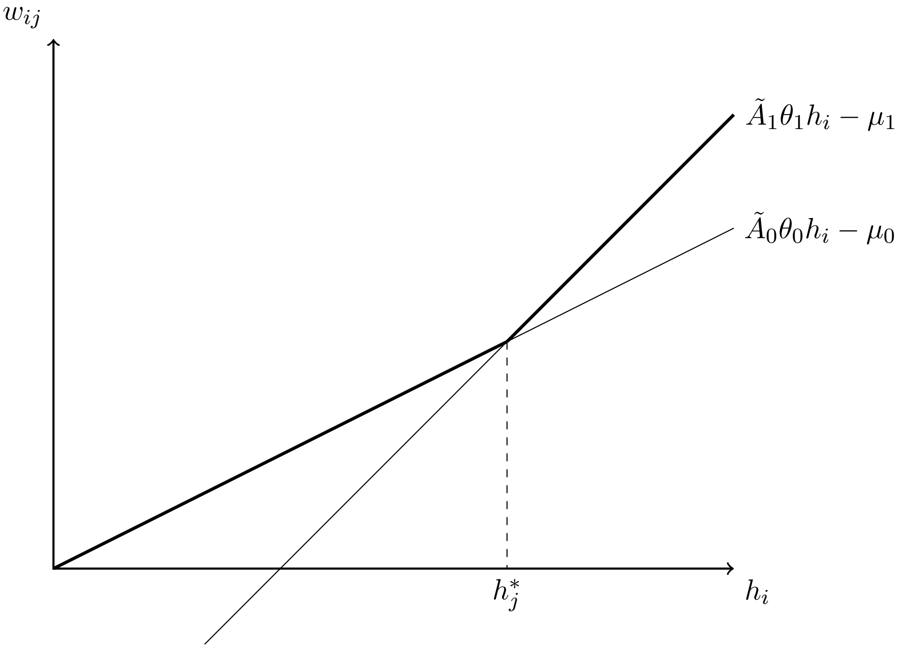
$$\pi_1^* = \frac{\theta_m h_i}{1+\theta_m h_i}.$$

Expected output

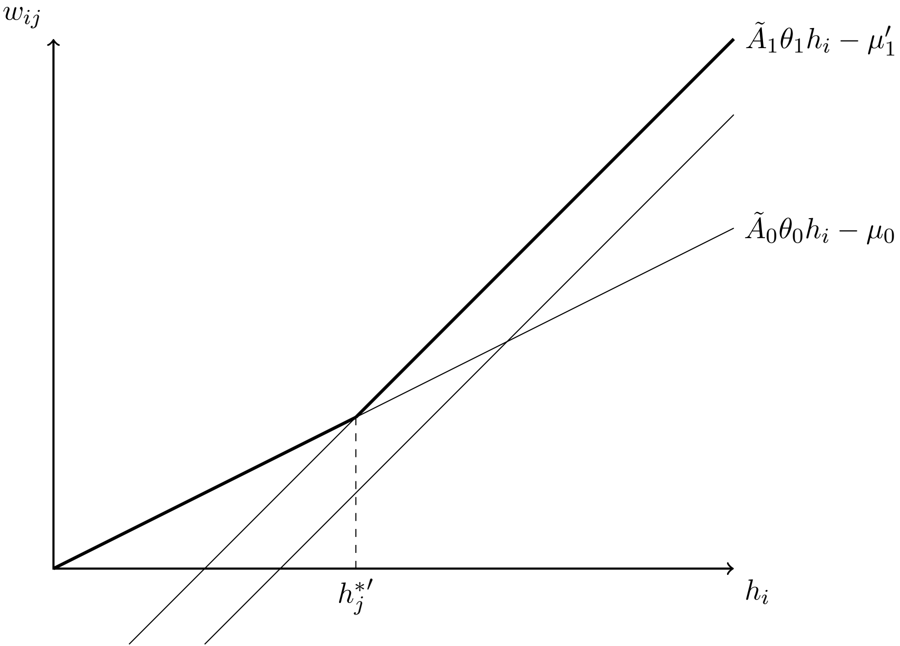
A worker type h operating k units of a machine type θ produces, in expectation,

$$F(A, k, \theta, h) = Ak\frac{\theta h}{1 + \theta h} \tag{1}$$

Machine assignment and wage setting by worker skill

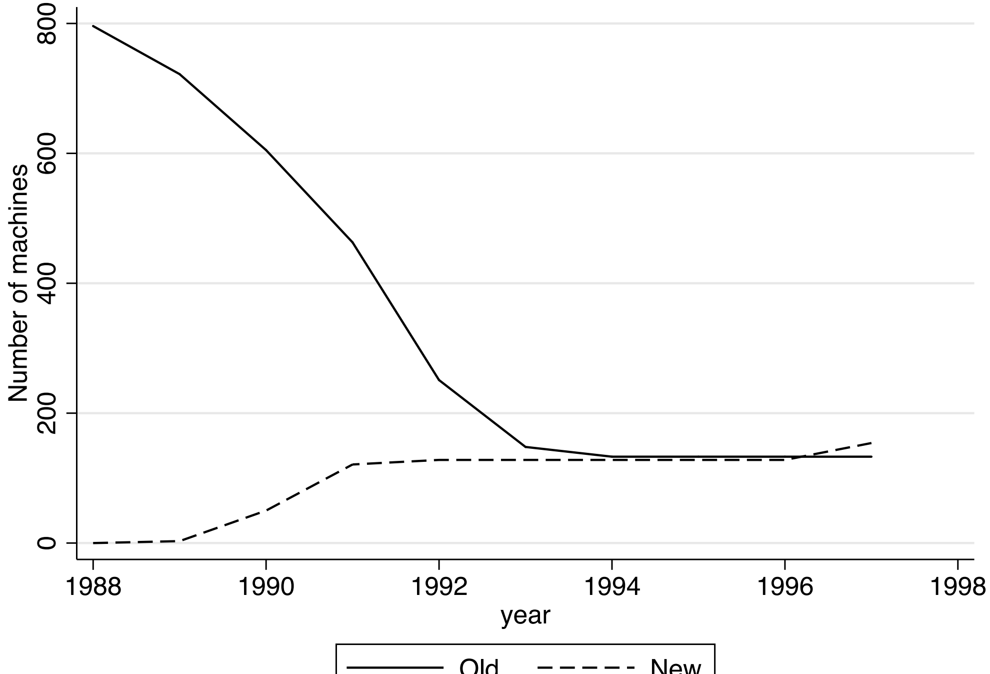


Technology upgrading by worker skill



A case study of a weaving mill

The number of old and new machines, 1988–1997

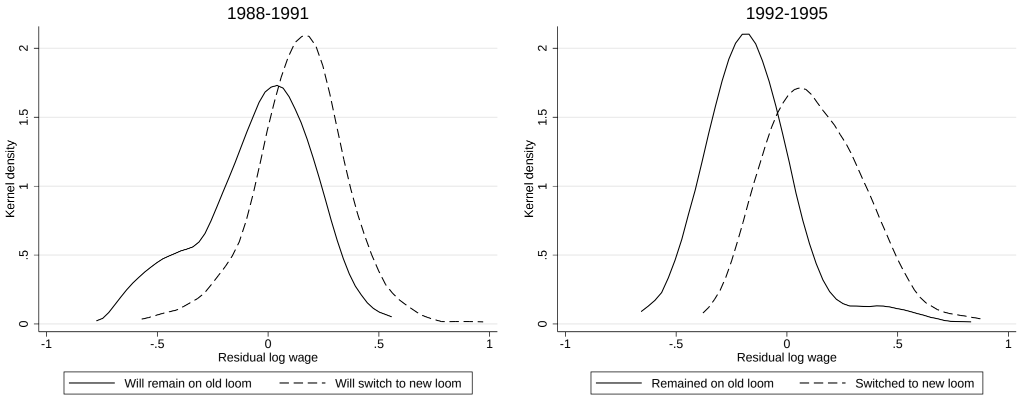


Differences between new and old machines—Regression estimates, 1991–1997

Dependent variable	Mean difference	Mean dep. var.	St. dev. dep. var.
Output (log)	0.820^{***}	5.49	0.475
Potential output (log)	0.790^{***}	5.94	0.449
Potential output/worker (log)	0.811^{***}	3.52	0.845
Output/potential output (log)	0.031^{*}	4.15	0.150
Percent downtime due to			
—scheduled maintenance	-3.20^{***}	2.73	3.30
—troubleshooting	-1.68^{***}	2.22	1.58
—change of warp	1.54^{**}	8.33	5.97
—change of weft	0.940^{***}	2.94	2.99
—other reasons	1.08	4.02	6.90
Total downtime	-0.961	20.38	9.74
Machine/worker	-2.64^{***}	11.32	2.29
Interventions/hour	-1.64	45.26	9.46

Notes: Number of observations: 341 machine-months observed between May 1991 and August 1997. Estimation: OLS with robust standard errors. For an accounting of how the estimation sample was constructed see Appendix B. In each equation, the dependent variable is regressed on a dummy for

Wage distribution before and after the adoption of new looms



Estimated kernel density of residual log wages relative to year mean. Bandwidth = 0.1. Left panel includes workers between 1988 and 1991 who do not yet work on a new loom (406 worker-years). Right panel includes workers between 1992 and 1995 (403 worker-years). Sample is limited to workers who appear in both time periods at least once.

Wage gain from moving from an old to a new machine

	(1) OLS	(2) Worker FE
New machine	0.167***	0.060***
	(0.021)	(0.020)
Age	0.075***	0.187***
	(0.007)	(0.021)
Age squared	-0.001***	-0.001***
	(0.000)	(0.000)
Number of observations	1,595	1,595
Number of workers	579	579
R^2	0.818	0.872

Notes: Dependent variable: log hourly wage. Sample: Person-years for continuing workers employed in the plant in 1989. Standard errors, clustered by worker, are reported in parantheses. Coefficients signifiantly different from zero at 1, 5 and 10 percent are marked by ***, ** and *, respectively.

The effect of machine type and worker quality on log output per machine

	(1) Production functic
Log number of weavers	0.109***
	(0.029)
New machine	-0.858**
	(0.335)
Log residual wage (as of 1989) of workers at the machine type	-9.91**
	(4.62)
New machine × log residual wage	38.53***
	(7.55)
Number of observations	261
R^2	0.733
Effect of the new machine at 25th percentile of the 1989 residual wage	-0.747
Effect of the new machine at 50th percentile of the 1989 residual wage	1.20
Effect of the new machine at 75th percentile of the 1989 residual wage	1.47

Notes: Dependent variable: log output per machine. Sample: machine-months for five types of loom. Estimat

Imported machines and wages in Hungary

The estimation sample over time

Year	Workers	Firms	Fraction importing (percent)	Import exposure (percent)
1992	10,853	1,823	35.27	16.57
1993	14,185	2,541	40.10	22.73
1994	14,695	2,773	39.07	27.26
1995	15,750	2,902	44.16	30.88
1996	15,419	2,775	48.74	34.35
1997	13,668	2,676	52.91	37.25
1998	15,239	2,754	55.22	40.04
1999	14,418	2,834	56.84	41.65
2000	14,805	2,966	55.89	43.44
2001	14,528	2,874	57.59	45.14
2002	15,907	2,345	53.40	45.99
2003	15,185	2,223	52.33	46.66
2004	15,261	2,281	49.86	46.66

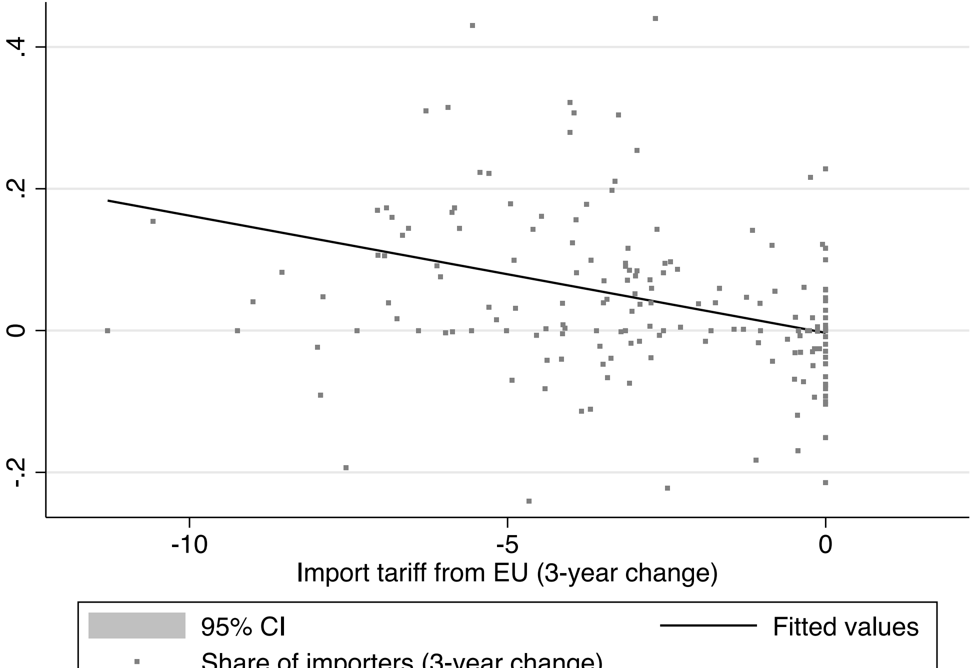
Notes: "Fraction importing" denotes the fraction of workers in the sample in importer occupations and importer firms ($\chi_{jot} = 1$). "Import exposure" is defined on a balanced sample of firm-occupations and denotes the same importer fraction in this balanced sample.

Average machinery tariffs

Year	Tariff on EU imports	Column 2 tariff
1992	9.40	9.70
1993	9.00	9.61
1994	8.69	9.61
1995	5.84	9.23
1996	3.18	9.02
1997	0.774	8.80
1998	0.572	8.56
1999	0.354	8.34
2000	0.176	8.33
2001	0.000	8.31
2002	0.000	8.33
2003	0.000	8.31

Notes: Table reports the unweighted average of tariffs on machinery imports from the European Economic Community (EU, second column), as well as the unweighted average of Column 2 tariffs on machinery (third column). Tariff rates are ad valorem percentages.

Occupations with faster tariff cuts adopt imported machines faster

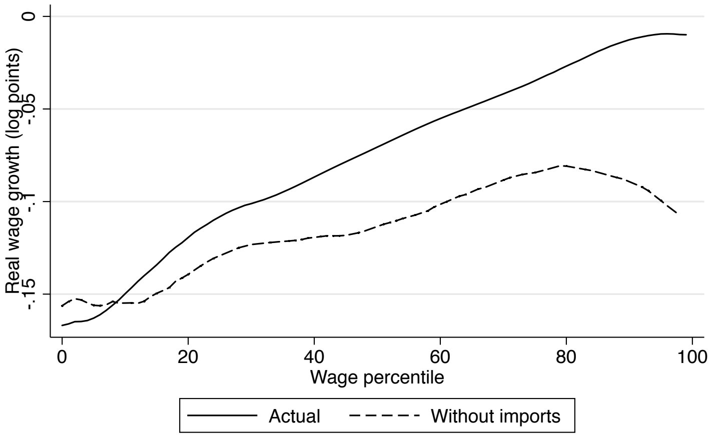


Wage inequality over time

Year	High-school premium	90/10 inequality
1992	0.168	0.978
1993	0.161	1.02
1994	0.178	1.01
1995	0.167	1.01
1996	0.180	1.06
1997	0.184	1.15
1998	0.184	1.16
1999	0.206	1.15
2000	0.205	1.17
2001	0.166	1.08
2002	0.189	0.959
2003	0.143	1.01
2004	0.179	1.03

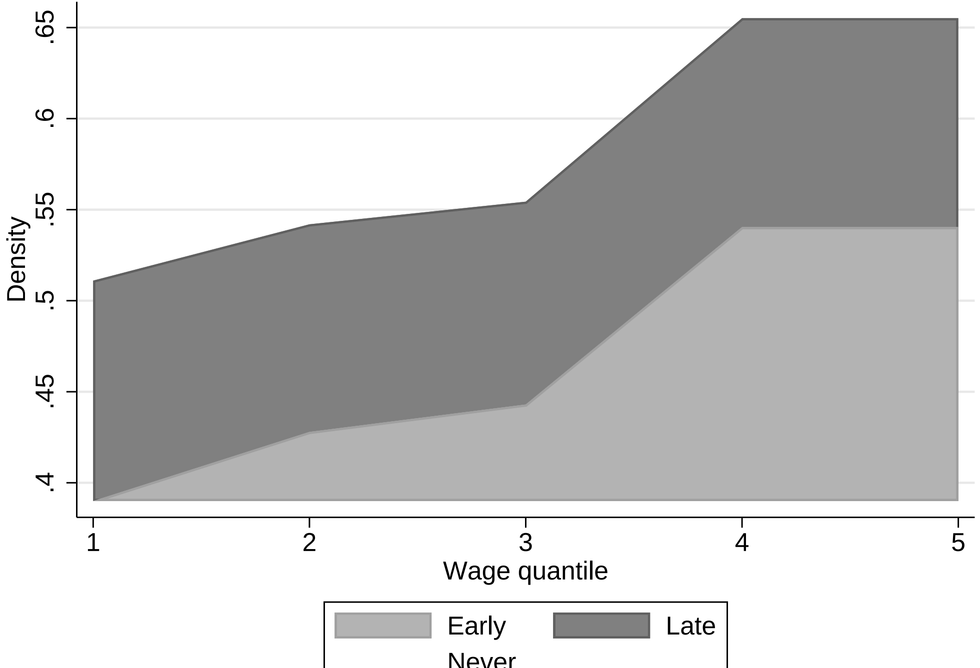
Notes: Table displays the wage gap between various groups of workers over time. The second column shows the wage difference (in log points) associated with a high-school degree (relative to primary school and vocational school), controlling worker gender, age and occupation. The third column shows the log point difference between

Actual and counterfactual wage change by wage percentile



Notes: Nonparametric estimates of log wage change between two periods by percentile of the wage distribution. Early period is 1992-1994 (15,205 worker-years), late period is 1998-2000 (17,475 worker-years). Firm-occupation cells that have already imported by 1994 are excluded. Counterfactua growth computed from firm-occupations cells that never import. Lowess curve with bandwidth of 0.33.

Among high-wage workers, early importers are overrepresented



Estimable equation

$$\ln w_{ij} \approx \ln(1-\beta)b + \frac{\beta}{(1-\beta)b}\left[\tilde{A}_0\theta_0 h_i - \mu_0 + \chi_{ij}(\tilde{A}_1\theta_1 - \tilde{A}_0\theta_0)(h_i - h_j^*)\right]. \quad \textbf{(??)}$$

We map this to the available data as follows.

$$\ln w_{ijot} = \alpha_{ot} + \nu_{jt} + \gamma_h h_i + \gamma_\chi \chi_{jot} + \gamma_{\chi h}\chi_{jot} h_i + u_{ijot}. \quad (2)$$

Creating a Bartik instrument

$$\hat{K}_{jot} = \frac{n_{jot} + n_{jo,t+1}}{\sum_l (n_{jlt} + n_{jl,t+1})} \times K_{jt}. \tag{3}$$

$$\hat{M}_{jot} = \frac{\hat{K}_{jot_0}}{\sum_l \hat{K}_{lot_0}} \times M_{ot}. \tag{4}$$

The effect of import exposure on wages

	(1) OLS	(2) OLS	(3) IV	(4) IV
Importer firm-occupation (dummy)	0.028^{***}	0.024^{***}	0.093^{**}	0.080^{*}
	(0.007)	(0.007)	(0.046)	(0.045)
High school diploma (dummy)	0.089^{***}	0.073^{***}	0.089^{***}	0.026^{**}
	(0.007)	(0.007)	(0.007)	(0.013)
High school diploma at importer firm-occupation (dummy)		0.027^{***}		0.105^{***}
		(0.009)		(0.026)
R^2	0.771	0.771	0.087	0.085
Number of observations	184,048	184,048	183,714	183,714
F-test for 1st stage				

Notes: The dependent variable is the log monthly earning of the worker in the given year. All specifications control for occupation-year and firm-year fixed effects. Worker controls include indicators for gender and schooling and a quadratic function of worker age. In columns 3 and 4, the importer dummy is instrumented by shift-share instruments, as explained in the main text. Standard errors, clustered by firm, are reported in parantheses. Coefficients significantly different from zero at 1, 5 and 10 percent are marked by ***, ** and *, respectively.

Robustness to various firm controls

	(1) No firm controls	(2) Capital stock	(3) Vintage	(4) Full control
Importer occupation	0.163***	0.084***	0.083***	0.049***
at importer firm (dummy)	(0.015)	(0.010)	(0.010)	(0.010)
Importer firm	0.170***	0.027***	0.027***	0.012
(dummy)	(0.012)	(0.010)	(0.010)	(0.010)
Book value of machinery		0.074***	0.074***	0.069***
(log)		(0.003)	(0.003)	(0.005)
Equipment bought 2–5			-0.054***	-0.042***
years ago (share)			(0.012)	(0.012)
Equipment bought 6 or			0.082**	0.067*
more years ago (share)			(0.032)	(0.037)
Firm is foreign owned				0.153***
(dummy)				(0.013)
R^2	0.381	0.459	0.460	0.479
Number of observations	172,212	172,212	172,212	172,212

Appendix

Worker wages

Workers earn w in agriculture. Large L number of workers, few K machines.

Net output over agriculture

$$x - w(1 - x) = x(1 + w) - w$$

Equilibrium

$$K(1 - x) \leq L$$

Output per worker hour

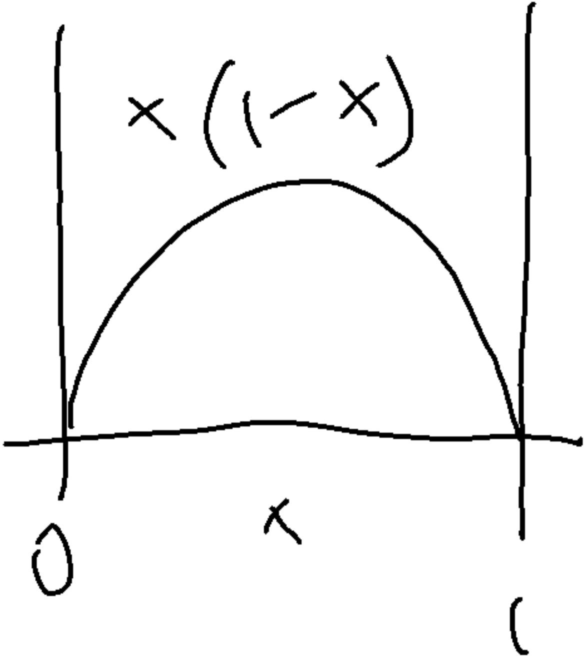
$$\frac{xK_i}{L_i} = \frac{x}{1-x}$$

Output per worker hour

Return to machine quality

$$\frac{\partial Q}{\partial \theta} = x(1-x)(1+w)$$

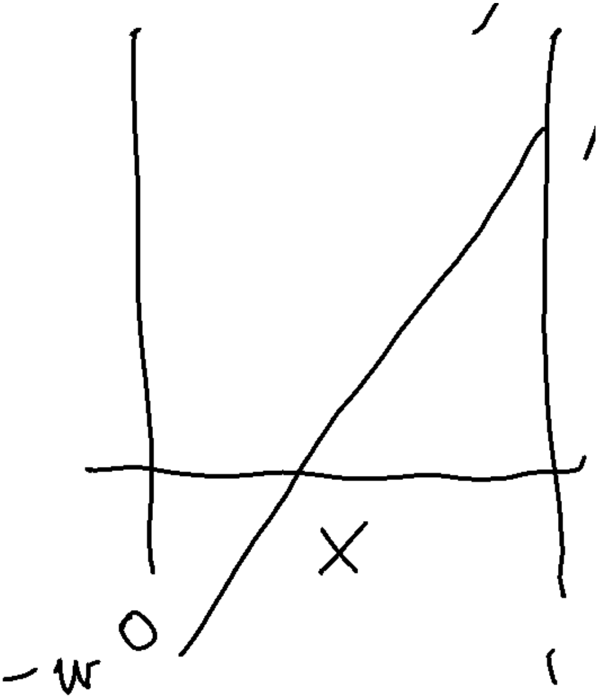
Return to machine quality



Value of a machine

$$Q_i - wL_i = x(1 + w) - w$$

Value of a machine



Innovation

Two types of innovation

1. Improve quality θ
2. Build more machines K

Suppose both cost the same amount of final goods.

Three epochs of innovation

1. Artisanal period. Labor is slack, wages are determined in agriculture. Improving machine quality has higher return. $x(1-x)(1+w) > x(1+w) - w$. x continues to increase.
2. Mass production. Value of a machine is high enough to produce more. K/L increases with constant x.
3. Automation. After all L has been absorbed from agriculture, wages start to rise. The returns to labor-saving machine quality improvement now exceed the value of an old machine. x keeps increasing.

Plus one

4. Singularity (never reached). As $x \to 1$, the ratio of machine time to worker time grows without bound. Nobody works, all work is done by robots. But to reach this state from a very large degree of automation (say, $x = 0.999$), labor has to capture almost all of the output, otherwise there is no incentive to innovate further.

Artisanal period

$$\frac{w}{1+w} \leq x \leq \sqrt{\frac{w}{1+w}}$$

Production function

Production function

$$Q_{it} = \Omega_{it}(\lambda K_{it}^{F} + K_{it}^{D})^{\alpha} L_{it}^{\beta} M_{it}^{\gamma}$$

with $\lambda > 1$

$$q_{it} \approx \omega_{it} + \alpha k_{it} + \beta l_{it} + \gamma m_{it} + \alpha(\lambda - 1)\frac{K_{it}^{F}}{K_{it}}$$

Patterns of capital imports

Data

- Hungarian Customs Statistics, 1992–2003
    - all *direct* exporter and importer
    - detailed by product (HS6): capital goods
    - and country of origin
- Balance Sheet and Earnings Statement
    - revenue, employment, material cost
    - capital: book value of equipment

Stocks and flows

- Imports are flows, equipment value is stock.
- Gross investment *flow*:

$$\hat{I}_{it} = K_{it} - (1 - \delta_{it}) K_{i,t-1}$$

  with $\hat{I}_{it} = \hat{I}^D_{it} + I^F_{it}$
- Imported equipment *stock*:

$$\hat{K}^F_{it} = (1 - \hat{\delta}_{it}) \hat{K}^F_{i,t-1} + I^F_{it}$$

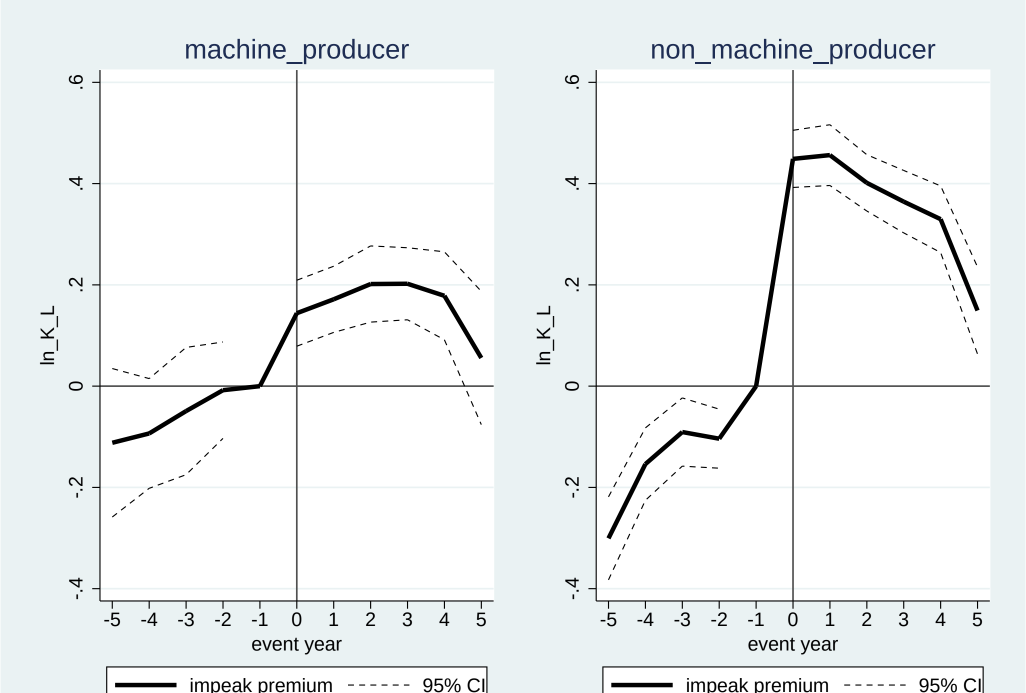
- Complications: what if $I^F_{it} > I_{it}$?

Distribution of investment rates (following Khan and Thomas, 2008)

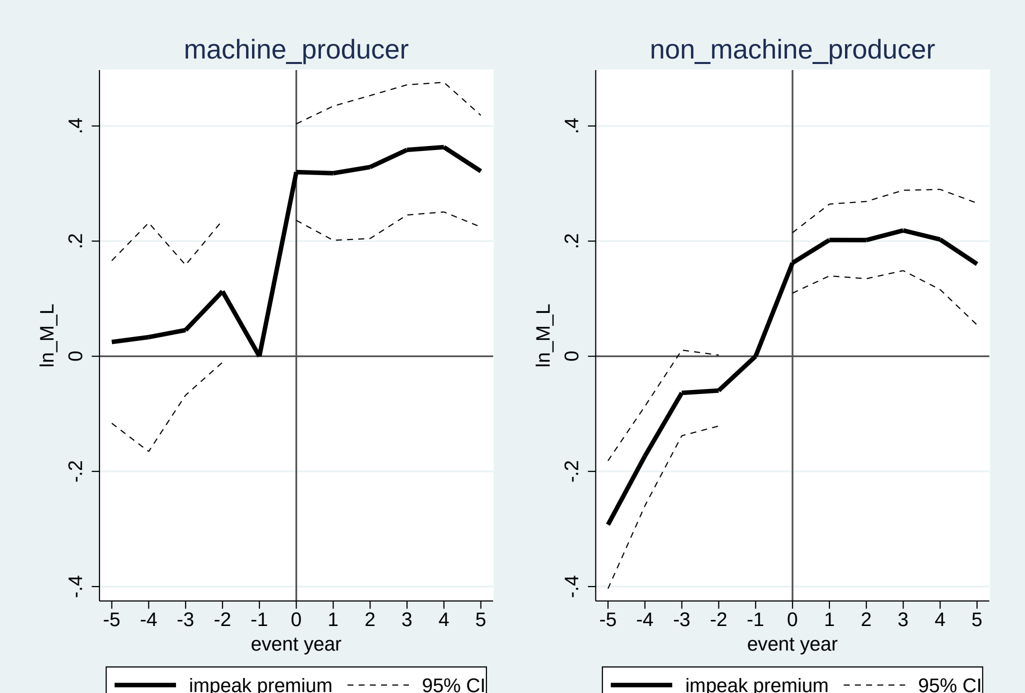
	Manufacturing 10+ employees	Non-machine manuf 10+ employees	Non-machine manuf all firm sizes
Average IR	0.321	0.270	-0.132
Average IR (winsor. 0.01)	0.378	0.335	0.338
Median IR	0.291	0.260	0.247
Inaction (%)	5.9	6.4	13.3
Positive investment (%)	85.9	85.0	77.0
Negative investment (%)	8.1	8.6	9.8
Positive spike (%)	59.9	56.9	54.1
Negative spike (%)	3.7	3.8	5.1
Observations	75,281	57,607	137,508

Notes: Inaction: abs(IR)<0.01, Positive spike: IR>0.2, Negative spike: IR<-0.2.
All samples exclude the first year of firms, where I_t equals K_t by construction.

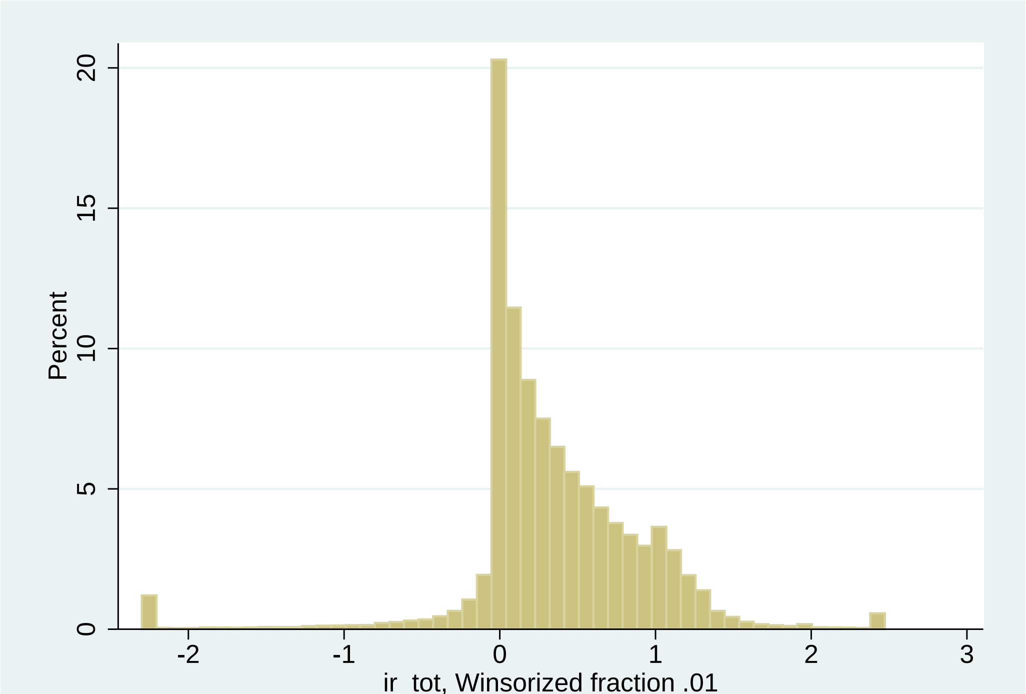
Capital intensity around import peaks



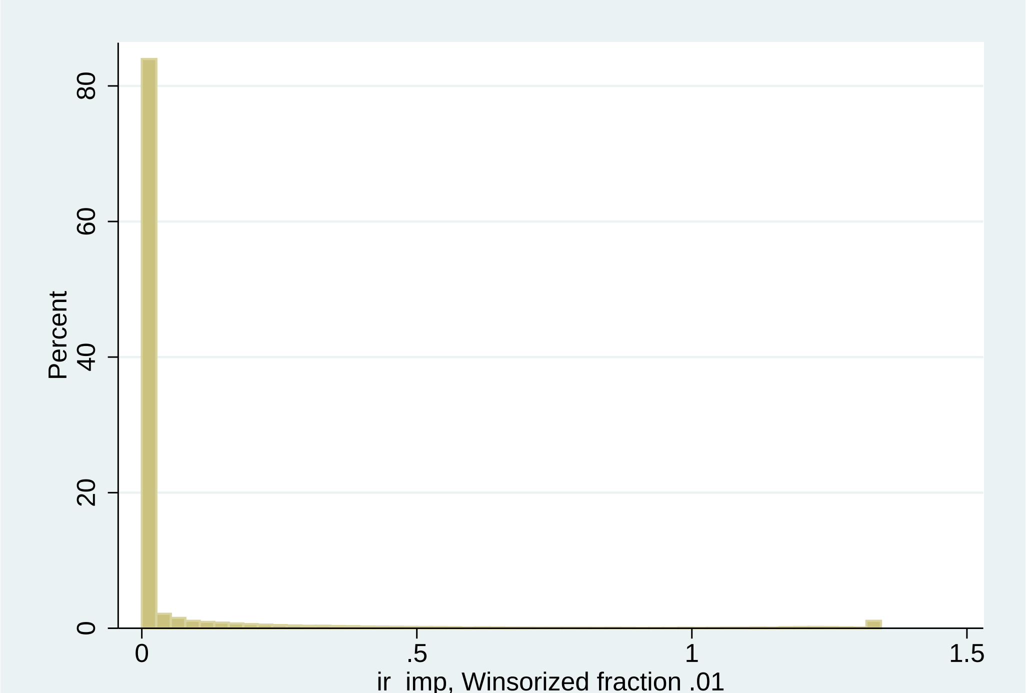
Material intensity around import peaks



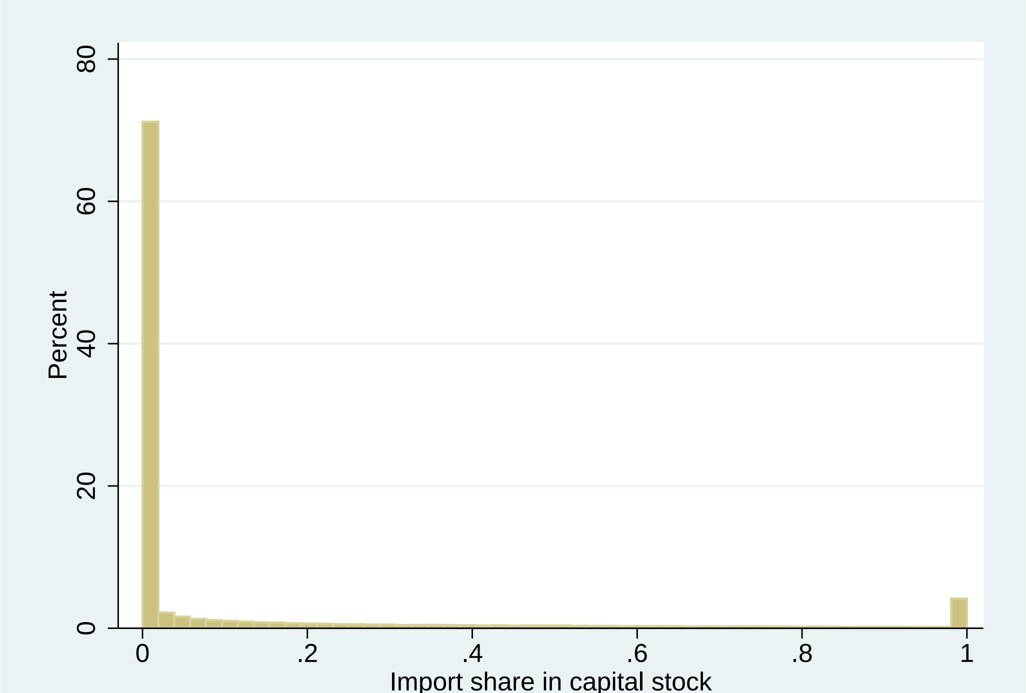
Investment rate distribution



Imported investment rate distribution

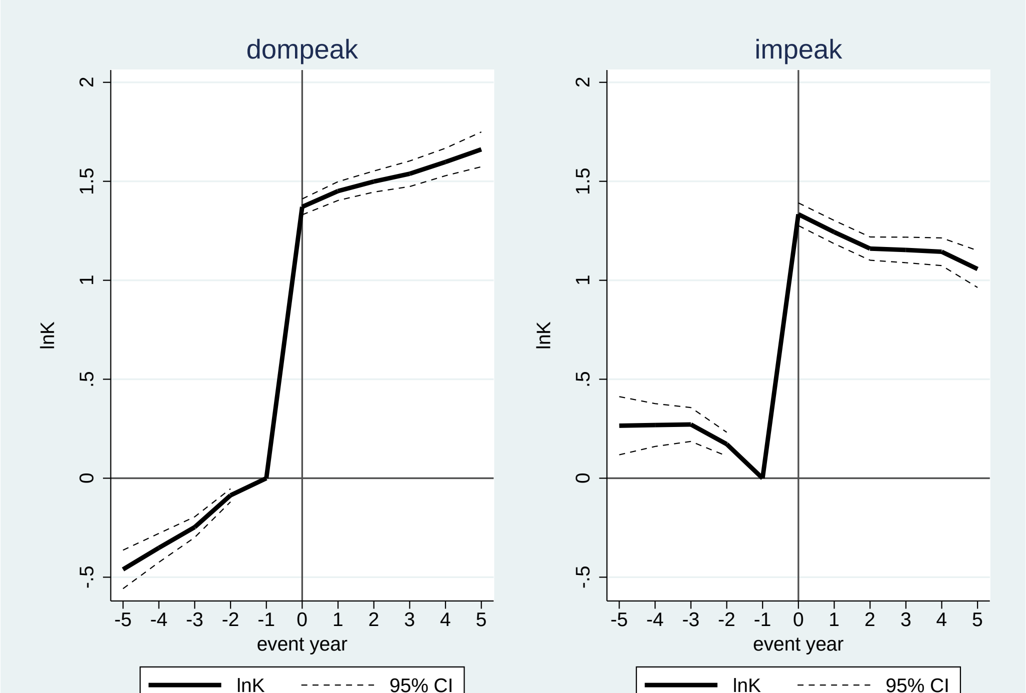


Import share in capital sock

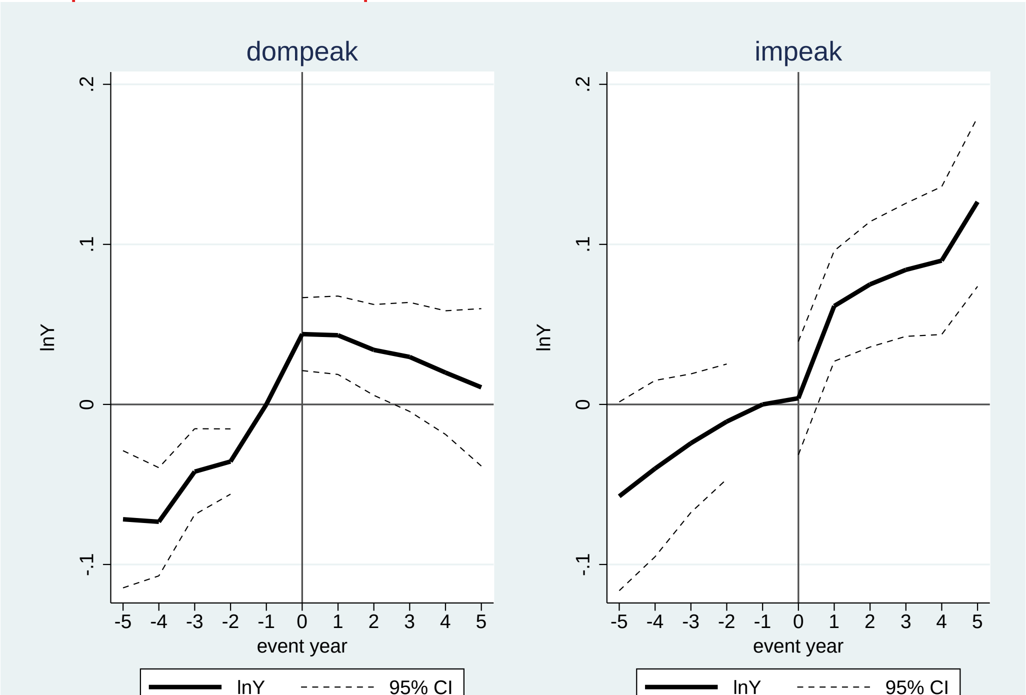


Event studies around large investments

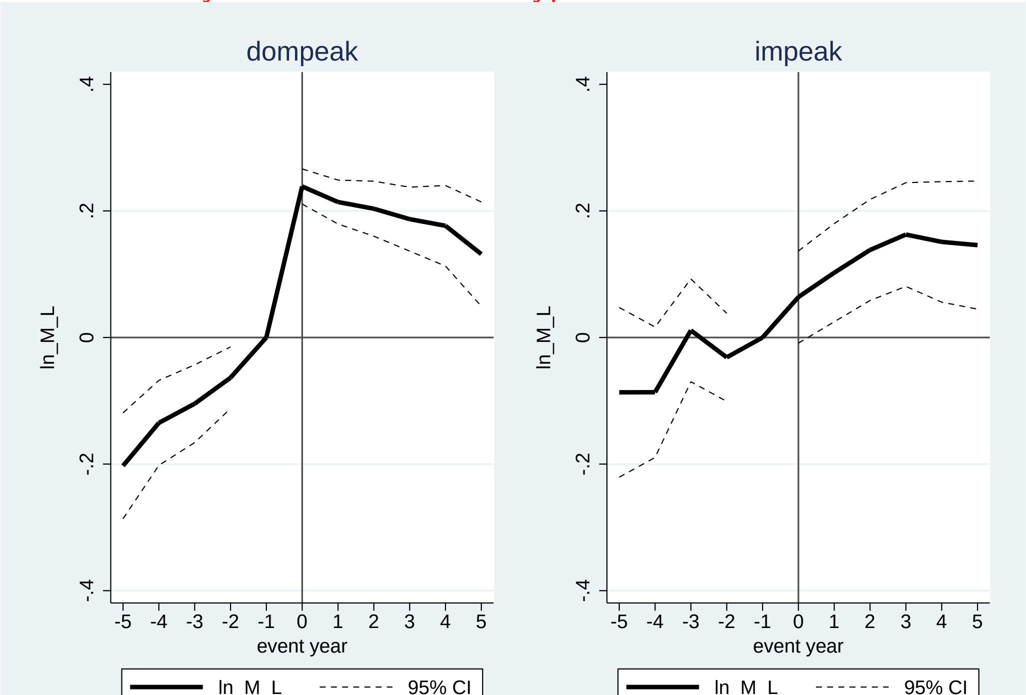
Capital stock increases by same amount (by construction)



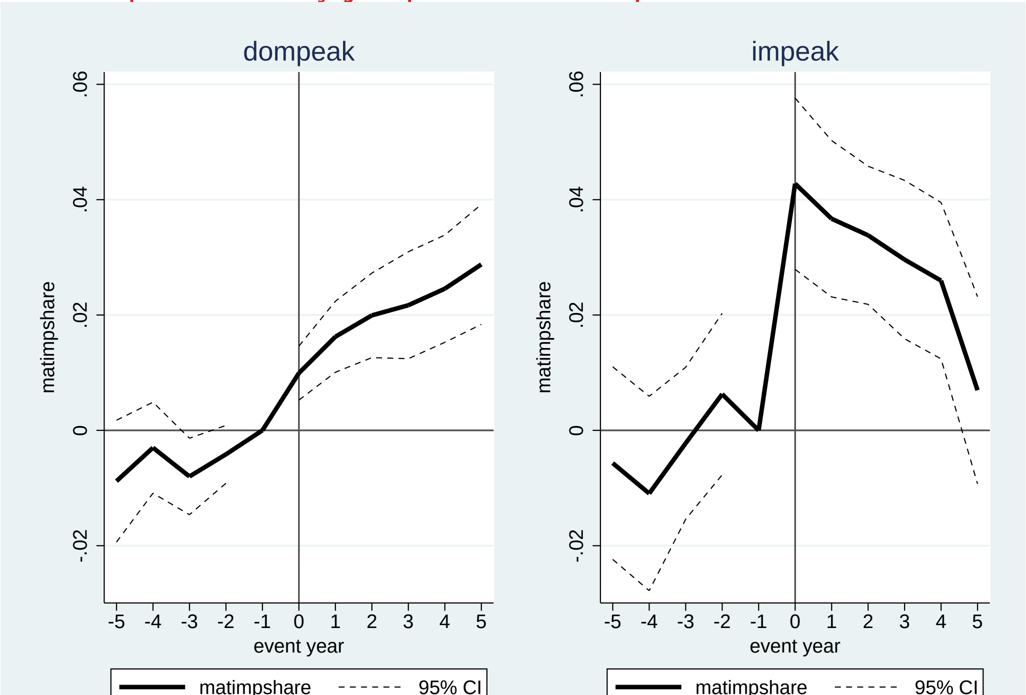
TFP improves more for imported investment



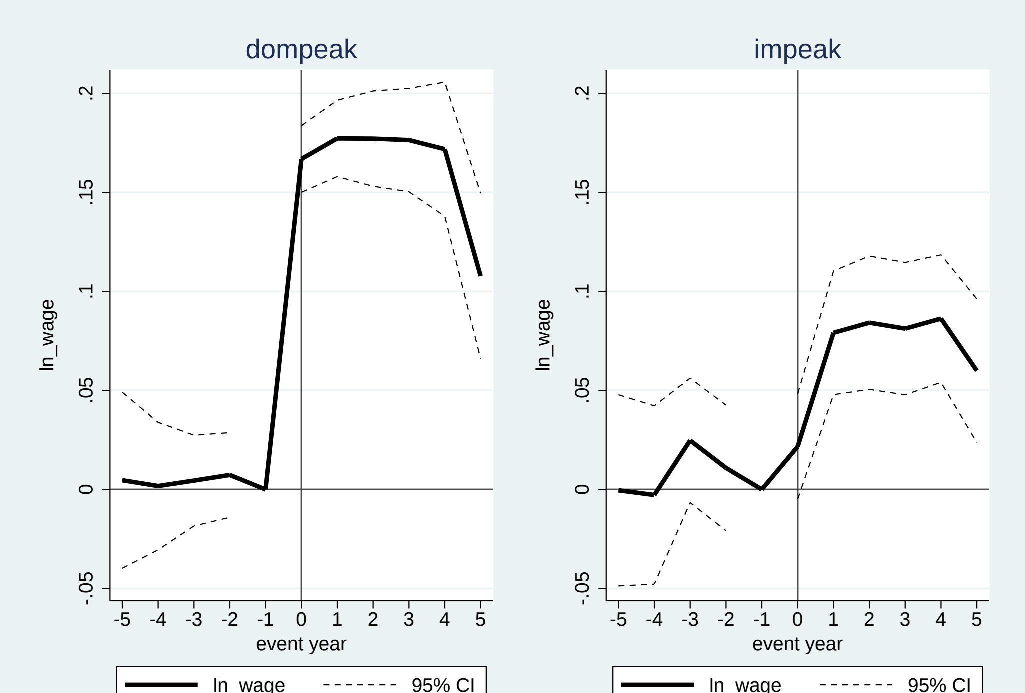
Material intensity increases for both types of investment



Material import intensity jumps more for imported investment

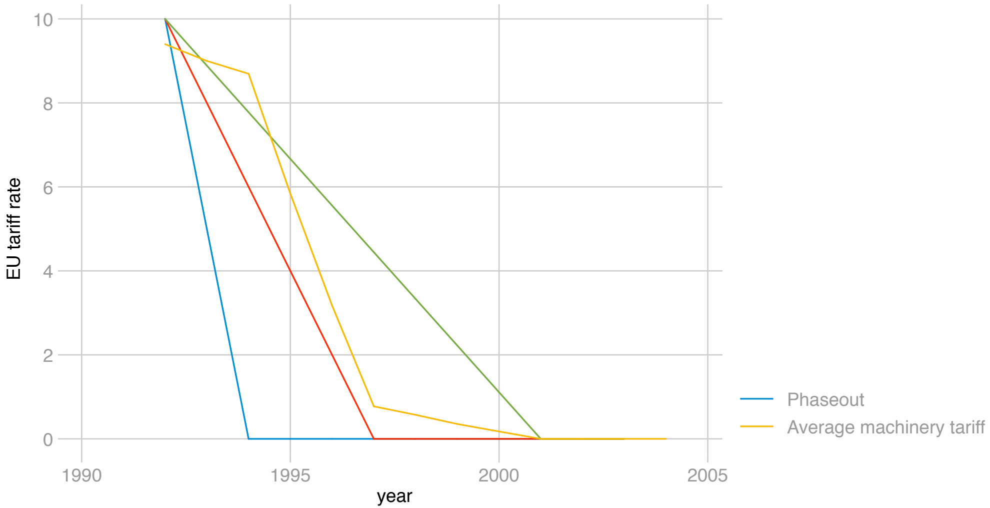


Average wage reacts to domestic investment

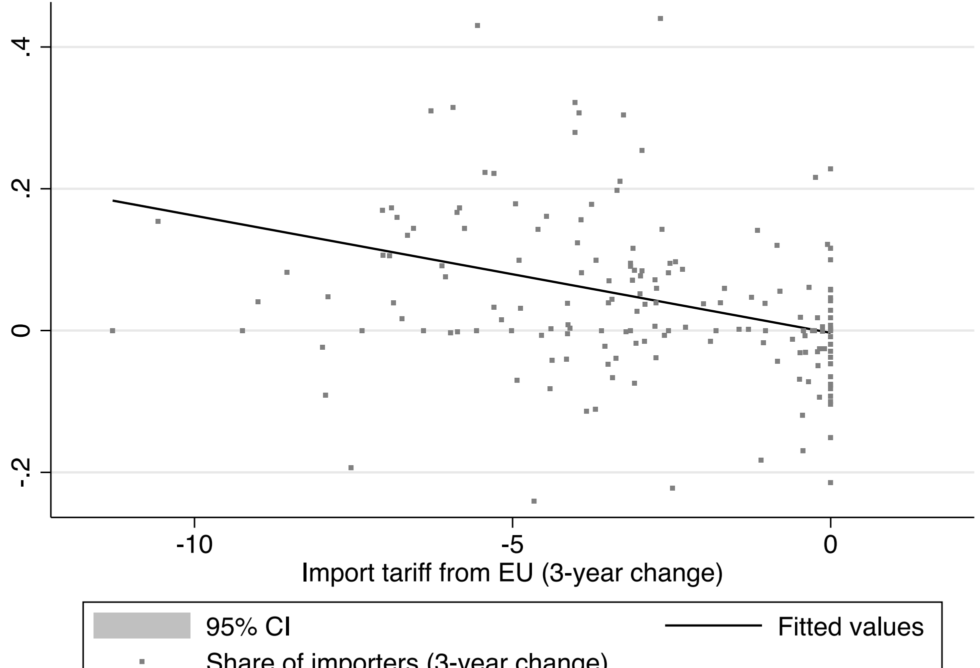


Identification

Interim Agreement with EEA (1991) phased out tariffs



Faster phaseout results in faster imports (Koren, Csillag and Köllő, 2019)



When do firms import?

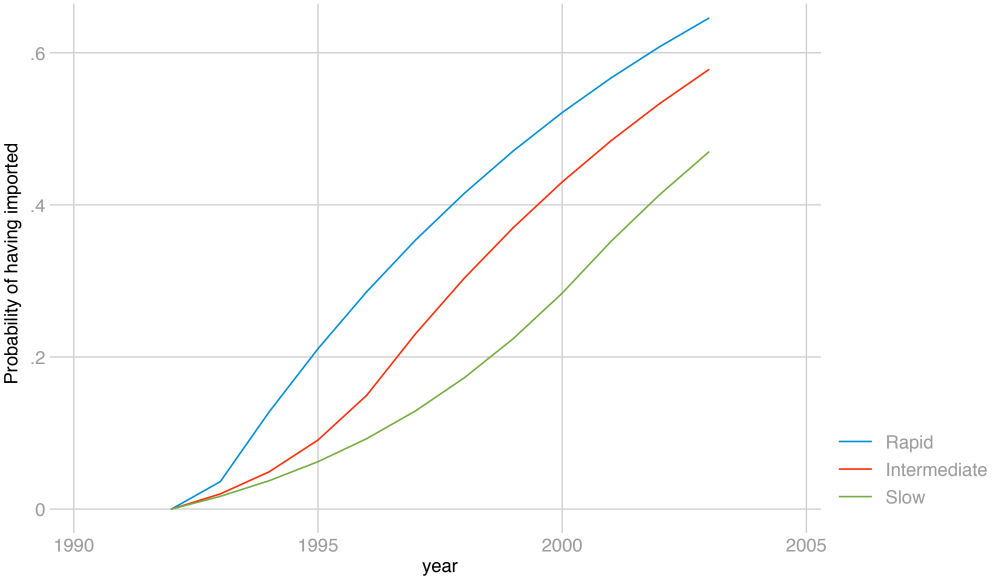
- Lumpy imported investment suggests fixed cost of importing (also see Halpern, Koren and Szeidl, 2015)
- Import if $p_t^F / p_t^D < f(L_{it})$.
- Hazard of *starting to import* (flow):

$$\Pr(K_{it}^F > 0 | K_{i,t-1}^F = 0) = \mu_{st} - \xi \Delta\tau_{st} L_{it}$$

- Probability of *having imported* in the past (stock):

$$\Pr(K_{it}^F > 0) \approx \tilde{\mu}_{st} - \xi L_{it-\mathsf{age}_{it}} \sum_{a=0}^{\mathsf{age}_{it}} \Delta\tau_{st-a}$$

Example of cumulated import hazards



Results

First stage

Depvar: having imported (dummy)	Pooled	Firm FE
cdtariffeu X size 0-10	-0.017***	0.009*
	(0.001)	(0.005)
cdtariffeu X size 10-50	-0.026***	-0.001
	(0.001)	(0.005)
cdtariffeu X size 50+	-0.046***	-0.019***
	(0.002)	(0.005)
lnK	0.048***	0.027***
	(0.002)	(0.001)
lnM	0.018***	0.007***
	(0.001)	(0.001)
lnL	0.008**	0.018***
	(0.003)	(0.003)
foreign (dummy)	0.321***	0.149***
	(0.011)	(0.022)
size dummies	yes	yes
age dummies	yes	yes
industry x year effects	yes	
year effects		yes
Observations	102,516	102,516
R-squared	0.296	0.211
Number of id		17,736
F-test	239.1	91.74

Notes: Robust standard errors (clustered by industry) are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Productivity

Depvar: lnY	Pooled		Firm FE	
	OLS	IV	OLS	IV
having imported (dummy)	0.199***	0.263***	0.086***	0.781***
	(0.015)	(0.075)	(0.012)	(0.112)
lnK	0.132***	0.129***	0.092***	0.073***
	(0.005)	(0.006)	(0.004)	(0.005)
lnM	0.413***	0.412***	0.297***	0.292***
	(0.009)	(0.010)	(0.010)	(0.010)
lnL	0.299***	0.299***	0.364***	0.353***
	(0.010)	(0.010)	(0.010)	(0.010)
foreign (dummy)	0.161***	0.140***	0.091**	-0.033
	(0.023)	(0.034)	(0.043)	(0.047)
size dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
industry x year effects	yes	yes		
year effects			yes	yes
Observations	102,516	102,516	102,516	102,516
R-squared	0.771	0.771	0.545	0.503
Number of id			17,736	17,736

Notes: Robust standard errors (clustered by industry) are in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Complementarity

Complementarity

- Are imported machines complementary with other inputs?
- If so, can explain
  - large gaps
  - divergence
- Two ways to measure complementarity (Brynjolfsson and Milgrom, 2013):
  - performance: $f_{xy} > 0$
  - behavior: $\partial x / \partial y > 0$

Positive cross derivative of output (Koren, Csillag and Köllő, 2019)

Table 4: The effect of machine type and worker quality on log output per machine

	(1)
	Production function
Log number of weavers	0.109***
	(0.029)
New machine	-0.858**
	(0.335)
Log residual wage (as of 1989) of workers at the machine type	-9.91**
	(4.62)
New machine × log residual wage	38.53***
	(7.55)
Number of observations	261
R^2	0.733
Effect of the new machine at 25th percentile of the 1989 residual wage	-0.747
Effect of the new machine at 50th percentile of the 1989 residual wage	1.20
Effect of the new machine at 75th percentile of the 1989 residual wage	1.47

Notes: Dependent variable: log output per machine. Sample: machine-months for five types of loom. Estimation: OLS. The average residual wage was measured by regressing individual log annual earnings (based on payment by results) in 1989 on age, age squared and type of machine fixed effects, and averaging the residual for workers employed at the given type of machine in the given month. Output is measured in million

Assortative assignment (Koren, Csillag and Köllő, 2019)

Table 2: The effect of worker quality on the probability that a worker was matched to a new machine

	(1) Machine-worker assignment
Log residual wage in 1989	2.63***
	(0.645)
Age	0.231**
	(0.100)
Age squared	-0.004***
	(0.001)
Tenure (years)	0.051***
	(0.015)
Number of observations	519
Pseudo-R^2	0.233
Standard deviation of the residual wage	0.128
Mean dependent variable	0.299

Notes: Dependent variable: 1 if the worker is assigned to a new machine, and 0 otherwise. Sample: person-years for continuing workers employed in the plant in 1989. Estimation: Probit. The residual wage was measured by regressing log payments by results in 1989 on age, age squared and type of machine fixed effects. Standard errors (in parentheses) are calculated from a 200-repetition bootstrap. Coefficients significantly different from zero at 1, 5 and 10 percent are marked by

Imported machines are more material intensive

Depvar: ln M/L	Pooled		Firm FE	
	OLS	IV	OLS	IV
having imported (dummy)	0.542***	0.706***	0.206***	1.218***
	(0.021)	(0.119)	(0.020)	(0.185)
foreign (dummy)	-0.032	-0.091*	0.109	-0.072
	(0.037)	(0.055)	(0.073)	(0.078)
size dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
industry x year effects	yes	yes		
year effects			yes	yes
Observations	102,516	102,516	102,516	102,516
R-squared	0.161	0.159	0.056	0.007
Number of id			17,736	17,736

Notes: Robust standard errors (clustered by industry) are in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Imported machines are more *imported* material intensive

Depvar: matimpshare	Pooled		Firm FE	
	OLS	IV	OLS	IV
having imported (dummy)	0.127***	0.110***	0.042***	0.148***
	(0.005)	(0.026)	(0.004)	(0.034)
foreign (dummy)	0.138***	0.144***	0.032**	0.014
	(0.009)	(0.013)	(0.014)	(0.015)
size dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
industry x year effects	yes	yes		
year effects			yes	yes
Observations	102,516	102,516	102,516	102,516
R-squared	0.186	0.186	0.010	-0.023
Number of id			17,736	17,736

Notes: Robust standard errors (clustered by industry) are in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Imported machines use higher quality labor

Depvar: ln wage	Pooled		Firm FE	
	OLS	IV	OLS	IV
having imported (dummy)	0.151***	0.586***	0.089***	0.796***
	(0.009)	(0.049)	(0.009)	(0.090)
foreign (dummy)	0.280***	0.125***	0.089**	-0.037
	(0.017)	(0.024)	(0.036)	(0.041)
size dummies	yes	yes	yes	yes
age dummies	yes	yes	yes	yes
industry x year effects	yes	yes		
year effects			yes	yes
Observations	102,516	102,516	102,516	102,516
R-squared	0.463	0.417	0.587	0.523
Number of id			17,736	17,736

Notes: Robust standard errors (clustered by industry) are in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$