THE GREAT DISPERSION IN ENERGY PRODUCTIVITY ACROSS FIRMS

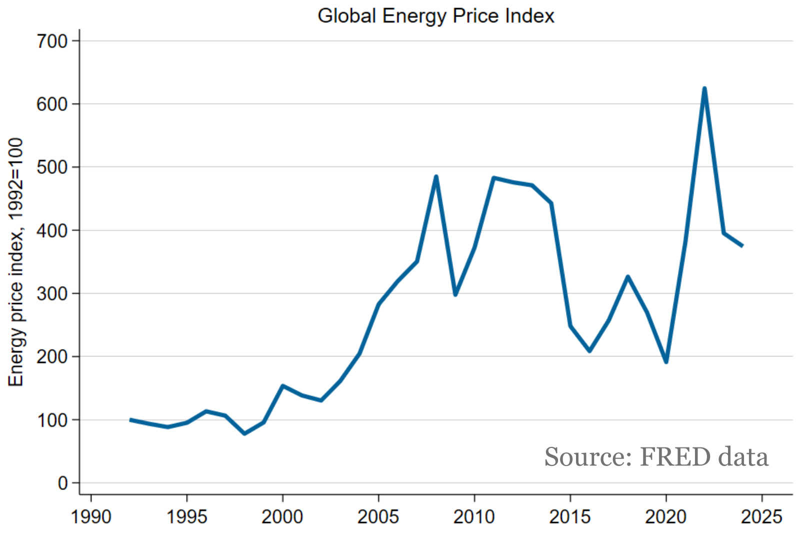
Josh De Lyon and Antoine Dechezleprêtre
OECD and CEP (LSE)

12 September 2025
NBER

BETTER POLICIES FOR BETTER LIVES

Increasingly important to improve energy productivity

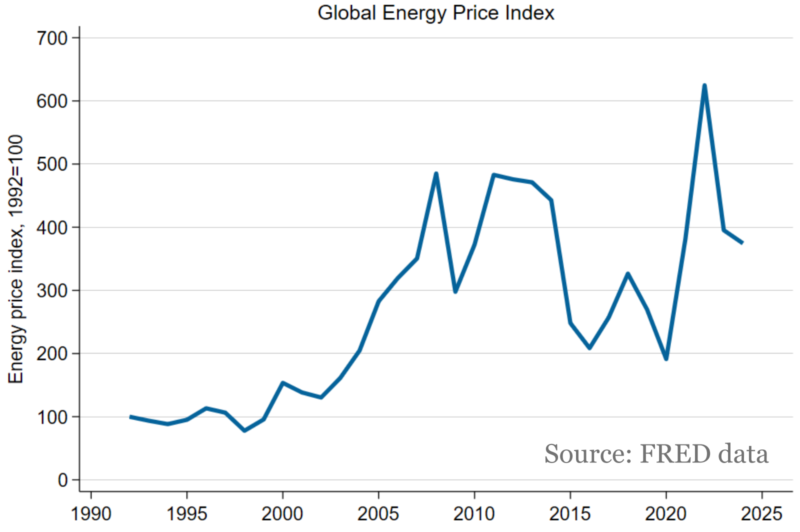
- Recent **energy price rises** reveal firms' vulnerability to energy costs



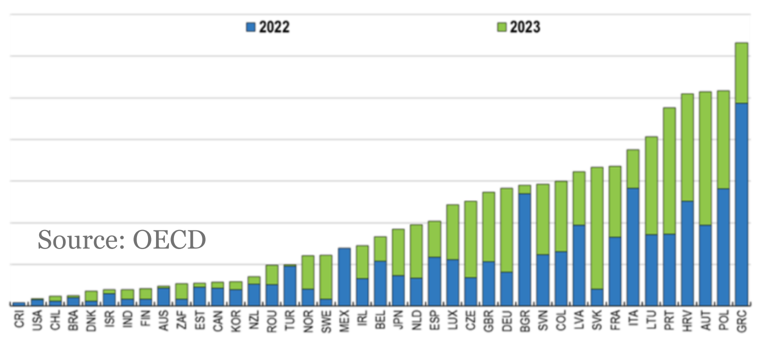
Increasingly important to improve energy productivity

- Recent **energy price rises** reveal firms' vulnerability to energy costs

- **Huge government spending** on energy support in 2022-23



Spending on energy support measures (% of GDP)

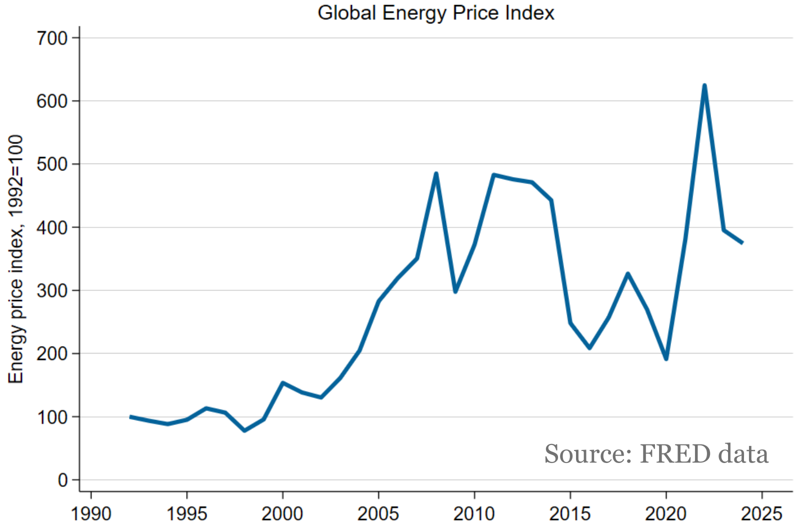


Increasingly important to improve energy productivity

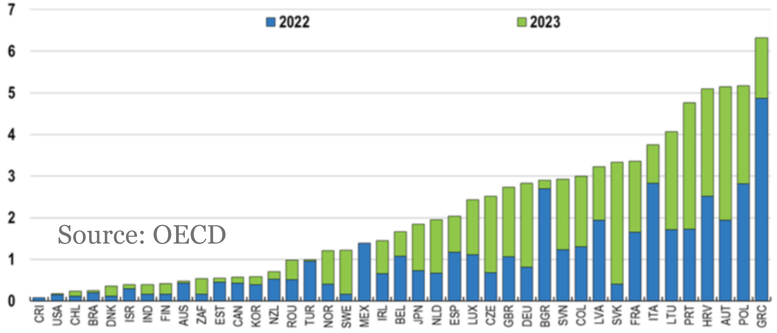
- Recent **energy price rises** reveal firms' vulnerability to energy costs

- **Huge government spending** on energy support in 2022-23

- Exposure to energy shocks is an issue for **national security**



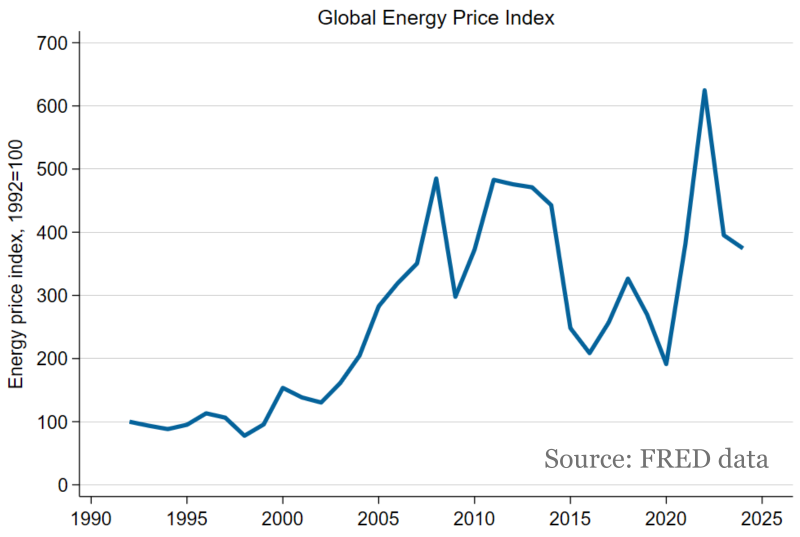
Spending on energy support measures (% of GDP)



Source: US Department of Energy

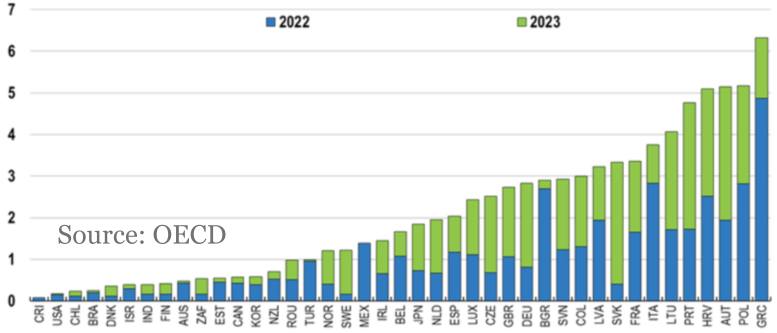
Increasingly important to improve energy productivity

- Recent **energy price rises** reveal firms' vulnerability to energy costs
- **Huge government spending** on energy support in 2022-23
- Exposure to energy shocks is an issue for **national security**
- **Energy efficiency must improve by 4% per year by 2030** to achieve net zero by 2050 (IEA)



Source: FRED data

Spending on energy support measures (% of GDP)



Source: OECD


U.S. DEPARTMENT of ENERGY

About · Our Mission · New Horizons · Topics · Consumer Savings

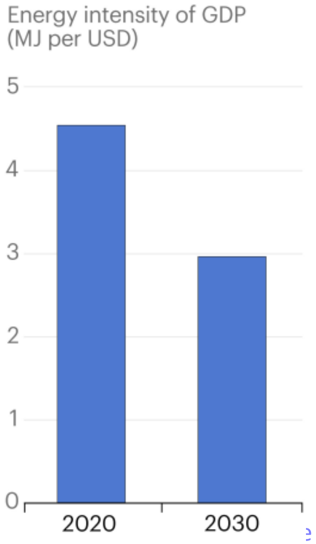
Office of Energy Efficiency & Renewable Energy › Resources › Why Energy Innovation Matters › Energy Independence and Se

Energy Independence and Security

Achieving U.S. energy independence would mean ending our nation's reliance on imported energy resources, securing our critical energy infrastructure against physical and cyber threats, and insulating our power system from market volatility and political instability abroad. Meeting these conditions will involve creating more American jobs in the power sector and related industries, such as manufacturing, and expanding America's energy supply chain so critical materials and components can be sourced domestically.

The U.S. Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy (EERE) are working to achieve U.S. energy independence and increase energy security by accelerating the growth of renewable energy sources.

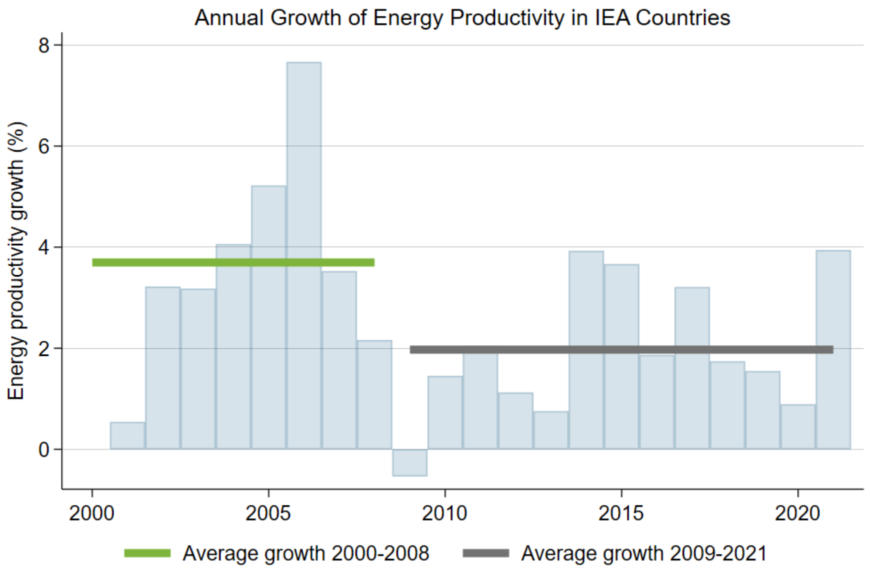
Source: US Department of Energy



Source: IEA

Energy productivity growth in manufacturing sector has slowed

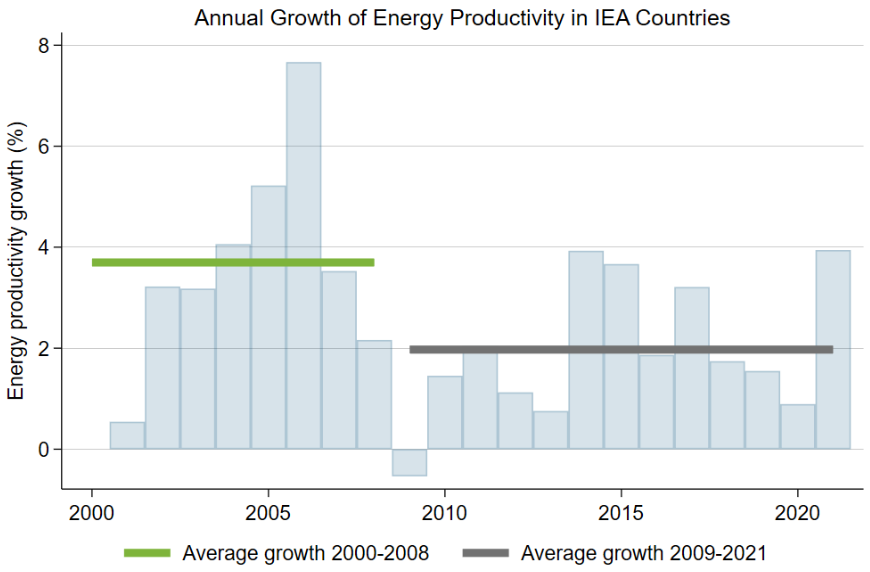
- **Slower energy productivity growth** in manufacturing sector since 2008



Source: IEA data

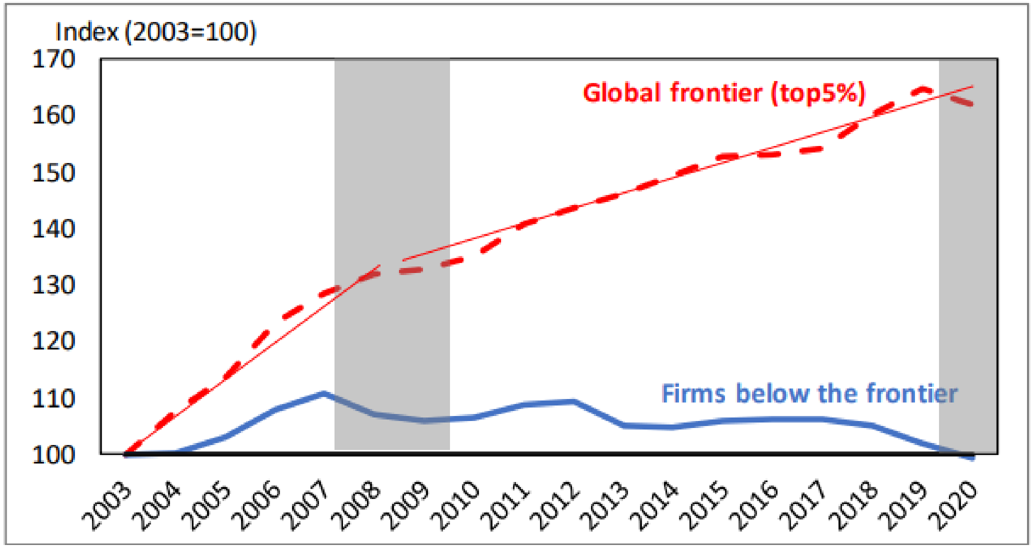
Energy productivity growth has slowed

- **Slower energy productivity growth** in manufacturing sector since 2008



Source: IEA data

- Large literature on slowdown in labour productivity growth particularly among laggards



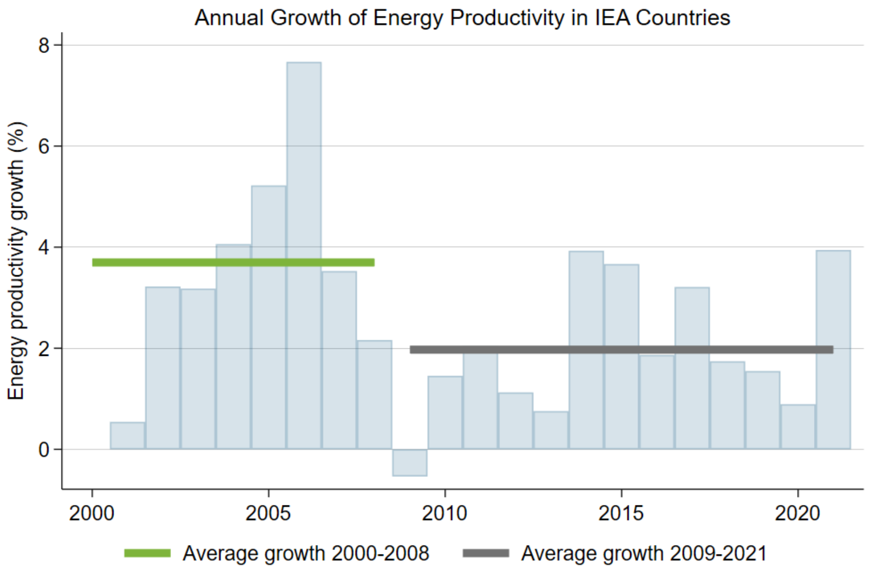
Source: André and Gal (2024)

Energy productivity growth has slowed

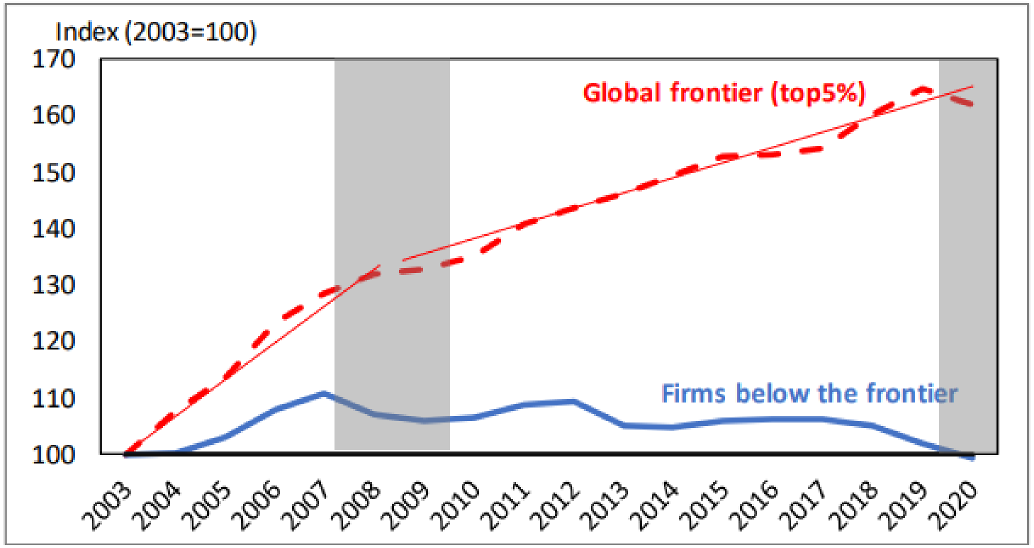
- Slower energy productivity growth in manufacturing sector since 2008

- Large literature on slowdown in labour productivity growth particularly among laggards

➢ But relatively little evidence on firm-level energy productivity



Source: IEA data



Source: André and Gal (2024)

Restricted

This paper: Evidence on firm-level energy productivity across countries

Research questions

1. How large are the differences between firms' energy productivity within industries?
2. What determines firms' energy productivity?
3. How important is it to address differences between firms' energy productivity?
4. Which policies can help improve energy productivity across the firm distribution?

This paper: Evidence on firm-level energy productivity across countries

Research questions

1. How large are the differences between firms' energy productivity within industries?
2. What determines firms' energy productivity?
3. How important is it to address differences between firms' energy productivity?
4. Which policies can help improve energy productivity across the firm distribution?

Data

- Centralised programs to use confidential and representative firm-level data for 8 countries
- Energy productivity (value added per energy quantity) and economic performance

This paper: Key findings

1. Huge dispersion between firms' energy productivity within industries
   – 90th percentile firm 21 times more energy productive than 10th percentile firm

2. Labour productivity is the most important predictor of energy productivity
   – Capital intensity, firm size, and age of capital also important

3. Large potential gains by improving energy productivity of least productive firms
   – Raising all firms to energy productivity of 25th percentile firm in their industry would reduce energy use by 45% to achieve same level of output

4. Policy-related factors associated with lower and declining dispersion
   – Higher energy prices, stronger competition and business dynamism, greater innovation activity, and access to credit associated with lower or declining dispersion

Related literature

- Firm heterogeneity in environmental performance
  - Lyubich, Shapiro, Walker (2018 AEA P&Ps) on US
  - Von Graevenitz and Rottner (2023 WP) & Petrick, Rehdanz, Wagner (2010 WP) on Germany; Klenow, Pasten, Ruane (2025 WP) on Chile
  - Wagner et al. (2020 WP) cross-country on emissions per employee

- Firm heterogeneity in economic performance
  - E.g. Berlingieri, Blanchenay, Criscuolo (2024 Research Policy); Andrews, Criscuolo, Gal (2019 WP); Berlingieri et al. (2020 IER); André and Gal (2024 WP); Autor et al. (2020 QJE); Syverson (2004 JPE, 2011 JEL); Hseih and Klenow (2009 QJE)

- Decomposition of manufacturing “clean-up”
  - Shapiro and Walker (2018 AER); Levinson (2015 JAERE); Rottner and Von Graevenitz (2024 Environmental & Resource Econ.); Murray Leclair (2025 WP)

DATA

Overview of data infrastructure

- Centralised programs for confidential datasets
- Possible input datasets:
  - Production data
  - Energy use survey
  - Business register
- Manufacturing sector, firms with employment >=20
- Representative of this sample at industry level (with survey weights)
- Cleaning and checks to harmonise across countries
- Data for Chile, Croatia, France, Indonesia, Lithuania, Netherlands, Portugal, Sweden
- Data covers 1995-2021, depending on country

Measurement of energy quantity

Two types of firm-level energy data:

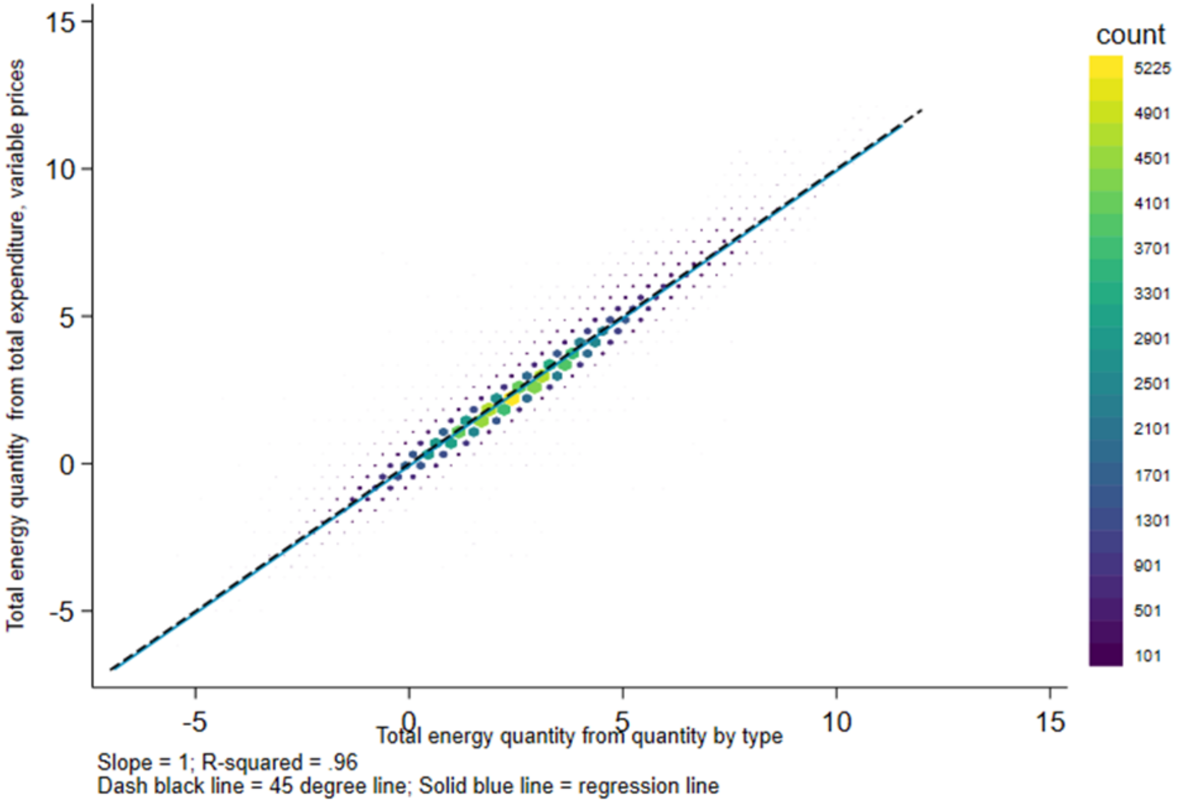
1. **Energy quantities by type of fuel**
   - Chile, France, Indonesia, Portugal, Sweden

2. **Total spending on energy**
   - Croatia, Lithuania, Netherlands
   - Estimate energy use using additional assumptions and data on:
     a. Energy type shares by industry (IEA, Eurostat)
     b. Energy prices at national level - can vary with expenditure of firm (IEA, Eurostat, Energiforsk)

Estimating energy quantity and emissions from total energy spending (Croatia, Lithuania, Netherlands)

- Algorithm to quantitatively estimate quantity of energy for each firm
- Accounts for:
  - Lower prices with greater quantity consumed
  - Different fuel type composition by sector
  - Lower electricity share with firms' total consumption

Comparison of energy quantity from total spending and true energy quantity for France



Final datasets for now

Firm-year level

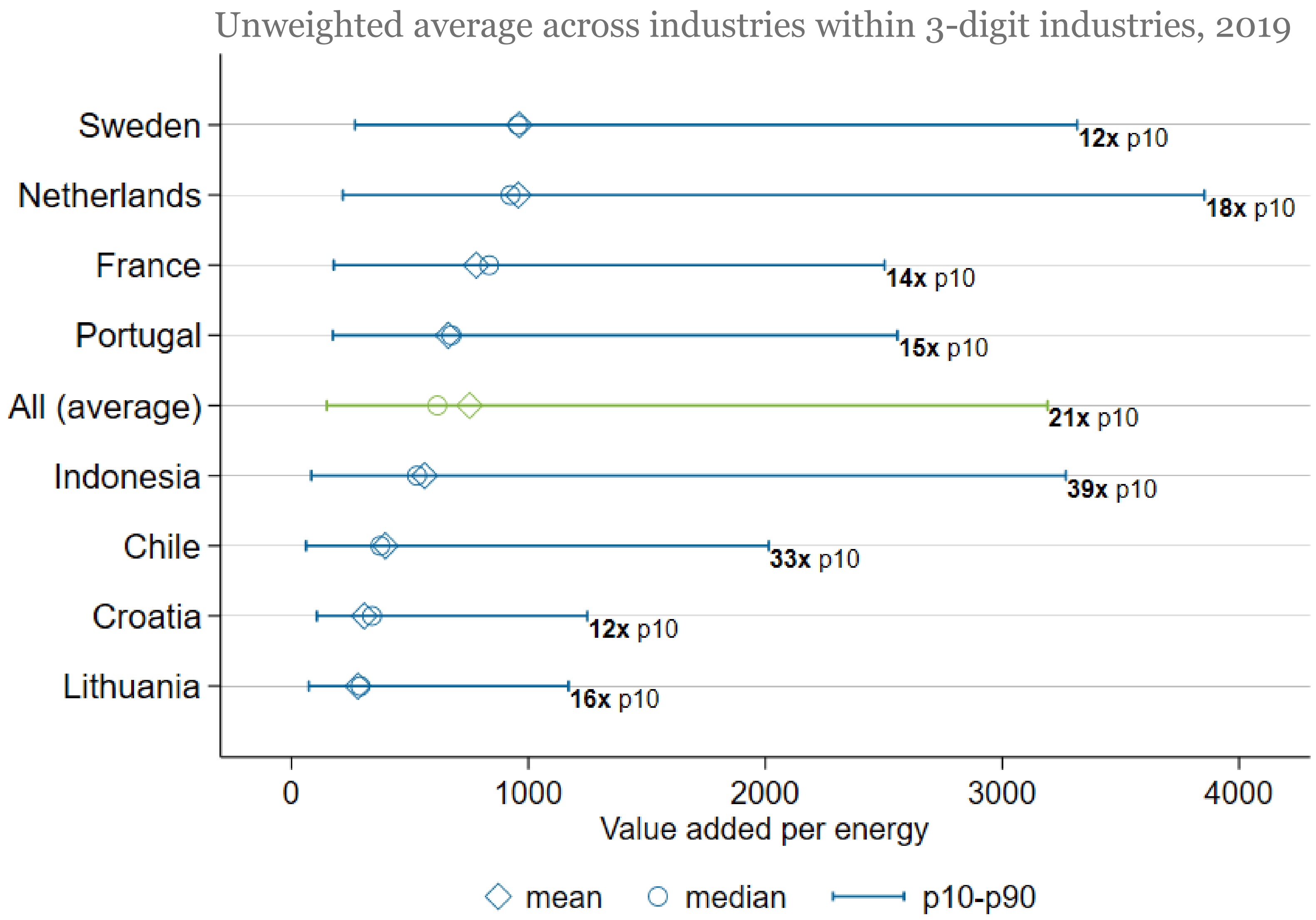
Country	Observations
Chile	61 798
Croatia	25 907
France	105 741
Indonesia	386 340
Lithuania	13 512
Netherlands	33 600
Portugal	52 647
Sweden	50 713

Country-industry-year level

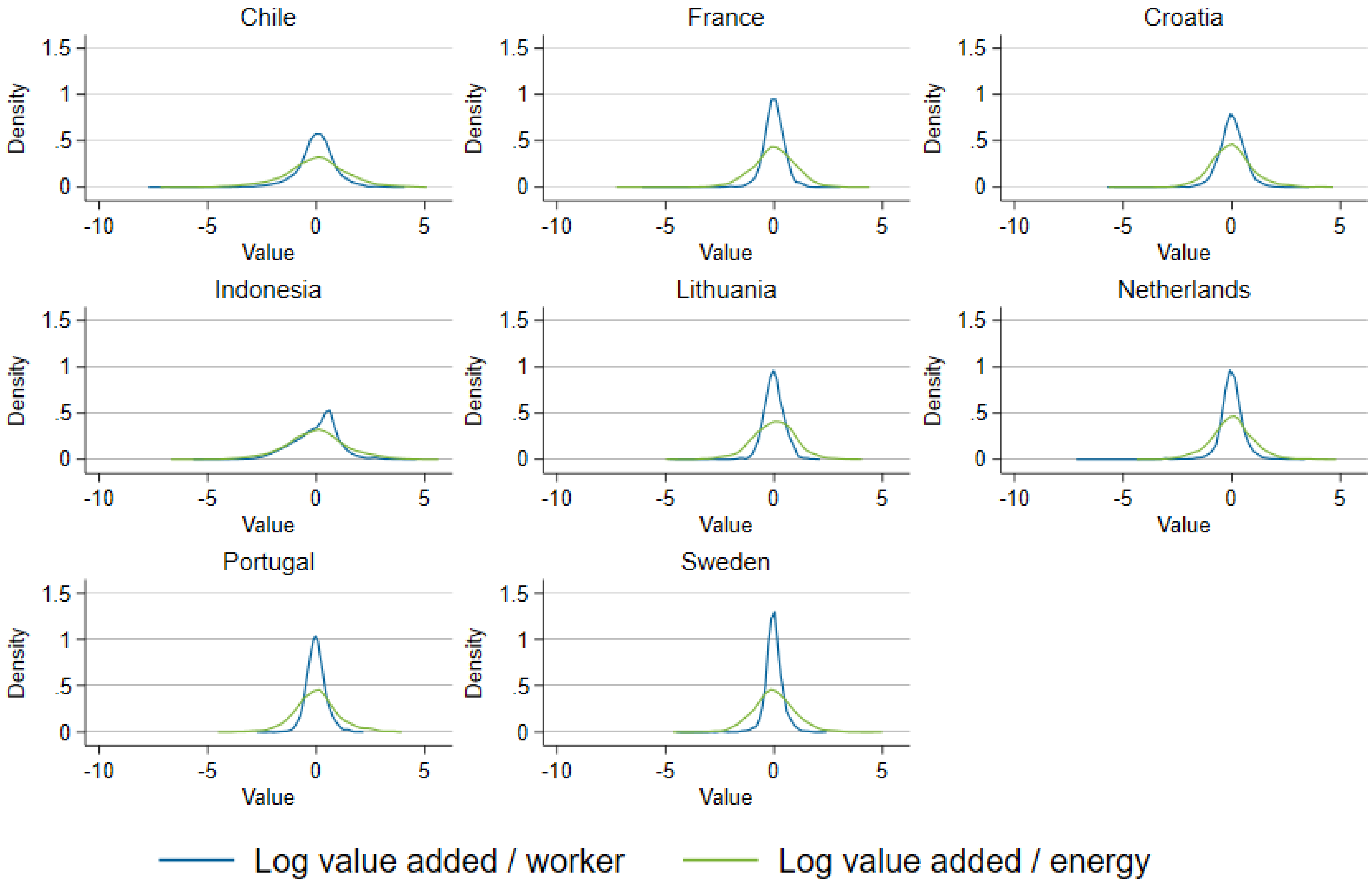
- 3 digit industry level
- Within-industry moments of the distribution
- Cross-country analysis
- Measure dispersion in cells with at least 10 firms
- 5 166 cells

1. DISPERSION IN ENERGY PRODUCTIVITY

90th percentile is 21 times more productive than 10th percentile firm



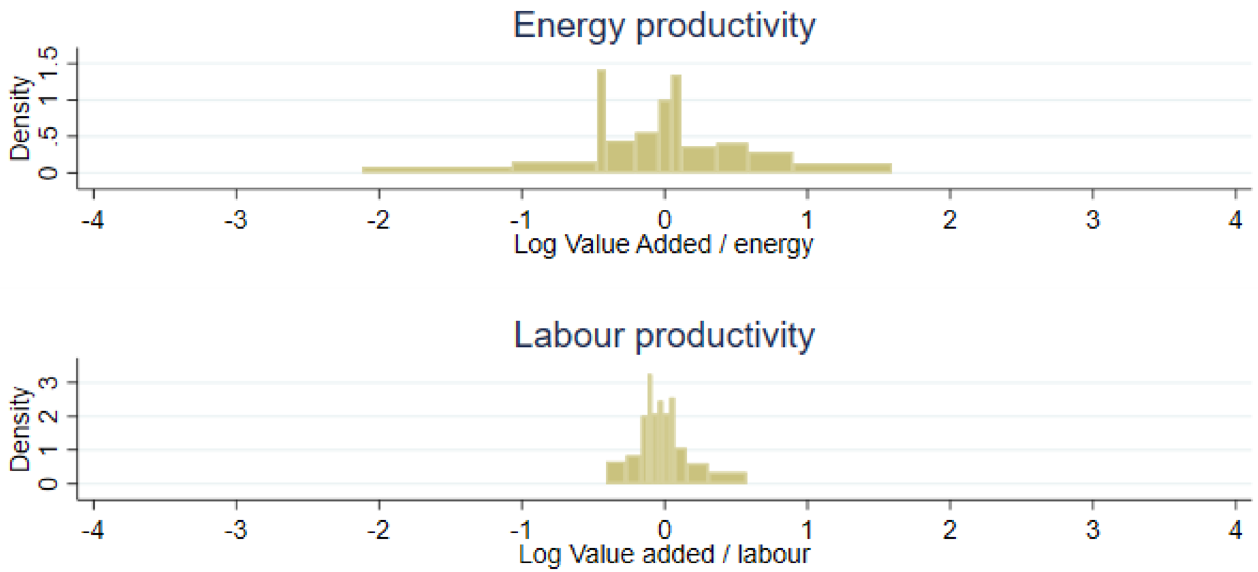
4 times more dispersion in energy than labour productivity



Within 4-digit industry distributions of energy and labour productivity by country (centred at 0)

Does dispersion hold for firms producing the same products?

- Data on sales by product in France
- 4 000 products in total
- In industry manufacture of basic iron and steel (2410), there are 123 products, including:
  - Remelting scrap ingots of iron or steel (24.10.14.20)
  - Ferrous products obtained by direct reduction of iron ore and other spongy ferrous products, (24.10.13.00)
  - Flat semi-finished products (slabs) (of stainless steel) (24.10.22.10)
  - Flat semi-finished products (of non-alloy steel) (24.10.21.10)
- Identify firms competing in the same product market
  - Defined as product with highest share of the firm's sales (and sales >50% or 80%)



Distribution of energy and labour productivity for firms producing only "Plastic doors, windows and their frames and thresholds for doors" (centred at 0)

Huge dispersion holds even within product markets

- In every case:
  - Huge dispersion in energy productivity
  - Far greater dispersion in energy than labour productivity

- This holds:
  1. Within 8 digit product markets

Product or industry-level	Dispersion in energy productivity	Dispersion in labour productivity
2 digit industry	17.5	3.2
3 digit industry	14	3.2
4 digit industry	12.2	3
4 digit industry (industry by size class sample)	13.1	3.1
4 digit industry by size class	12.1	3
4 digit industry (8 digit product sample)	9.4	2.8
6 digit product (highest firm share)	10.2	2.9
6 digit product (firm share >50%)	9.9	2.9
6 digit product (firm share >50%)	9.6	2.9
8 digit product (highest firm share)	8.9	2.8
8 digit product (firm share >50%)	8.7	2.8
8 digit product (firm share >80%)	8.6	2.9
8 digit product (all firms with proportionality assumption)	10	2.9
8 digit product (single product firms)	8.5	2.8

Table shows within product or industry dispersion in energy, and labour productivity (90th percentile / 10th percentile)

Huge dispersion holds even within product markets

- In every case:
  - ➢ Huge dispersion in energy productivity
  - ➢ Far greater dispersion in energy than labour productivity

➢ This holds:

1. Within 8 digit product markets
2. Within industry-size bands

Product or industry-level	Dispersion in energy productivity	Dispersion in labour productivity
2 digit industry	17.5	3.2
3 digit industry	14	3.2
4 digit industry	12.2	3
4 digit industry (industry by size class sample)	13.1	3.1
4 digit industry by size class	12.1	3
4 digit industry (8 digit product sample)	9.4	2.8
6 digit product (highest firm share)	10.2	2.9
6 digit product (firm share >50%)	9.9	2.9
6 digit product (firm share >50%)	9.6	2.9
8 digit product (highest firm share)	8.9	2.8
8 digit product (firm share >50%)	8.7	2.8
8 digit product (firm share >80%)	8.6	2.9
8 digit product (all firms with proportionality assumption)	10	2.9
8 digit product (single product firms)	8.5	2.8

Table shows within product or industry dispersion in energy, and labour productivity (90^{th} percentile / 10^{th} percentile)

Huge dispersion holds even within product markets

- In every case:
  - Huge dispersion in energy productivity
  - Far greater dispersion in energy than labour productivity

- This holds:
  1. Within 8 digit product markets
  2. Within industry-size bands
  3. Measuring output in quantity terms

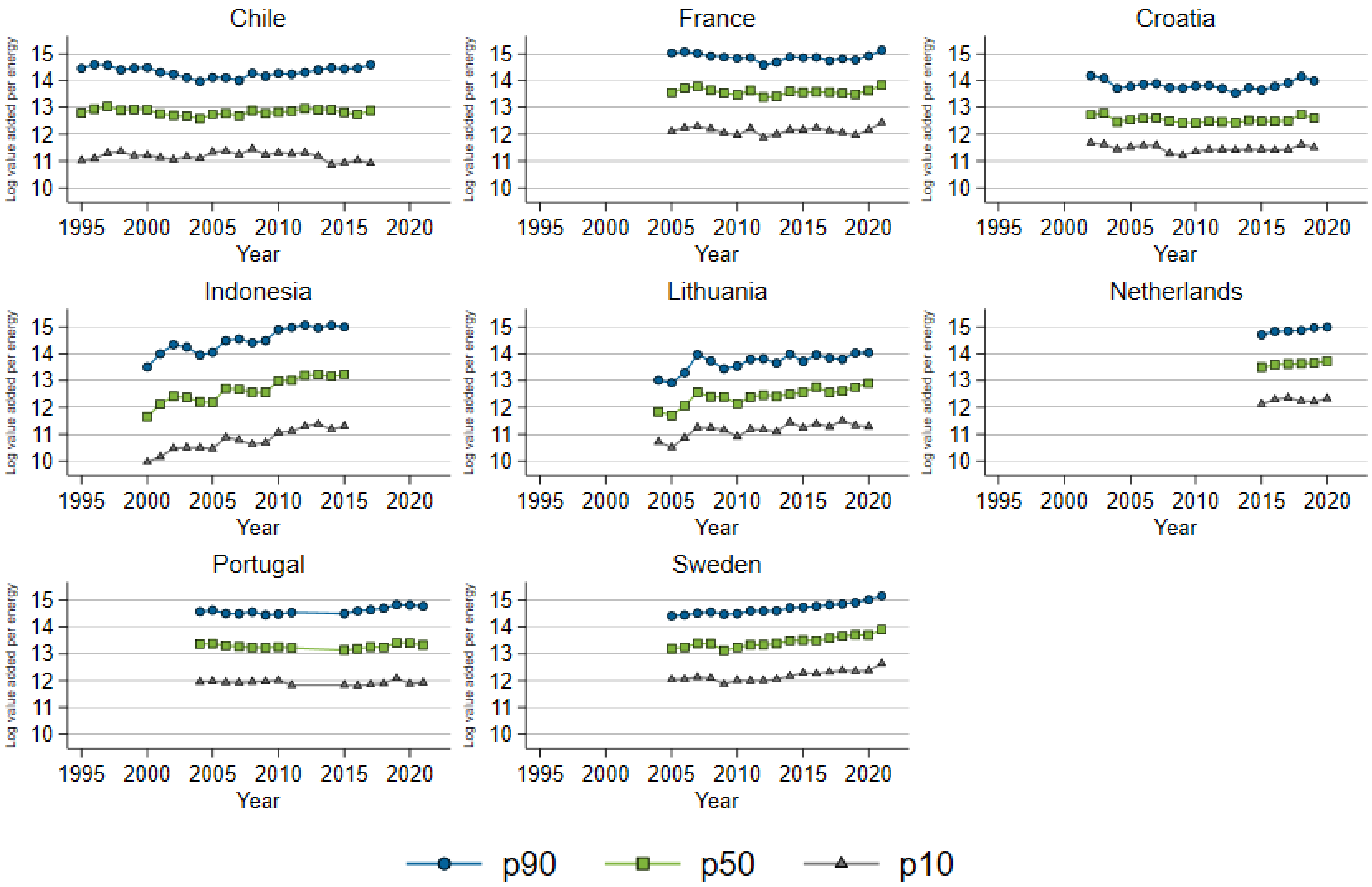
Why more dispersion in quantity?

Physical productivity negatively correlated with price

(Foster, Haltiwanger, Syverson, 2008)

Product or industry-level	Dispersion in energy productivity	Dispersion in labour productivity	Dispersion in energy productivity (quantity)	Dispersion in labour productivity (quantity)
8 digit product (all firms with proportionality assumption)	10	2.9	28.2	21.6
8 digit product (single product firms)	8.5	2.8	18.4	14.3

Dispersion is stable or increasing over time



Energy productivity at the 10^{th}, 50^{th}, and 90^{th} percentiles over time

Rate of catch-up in energy productivity is mostly stable or decreasing

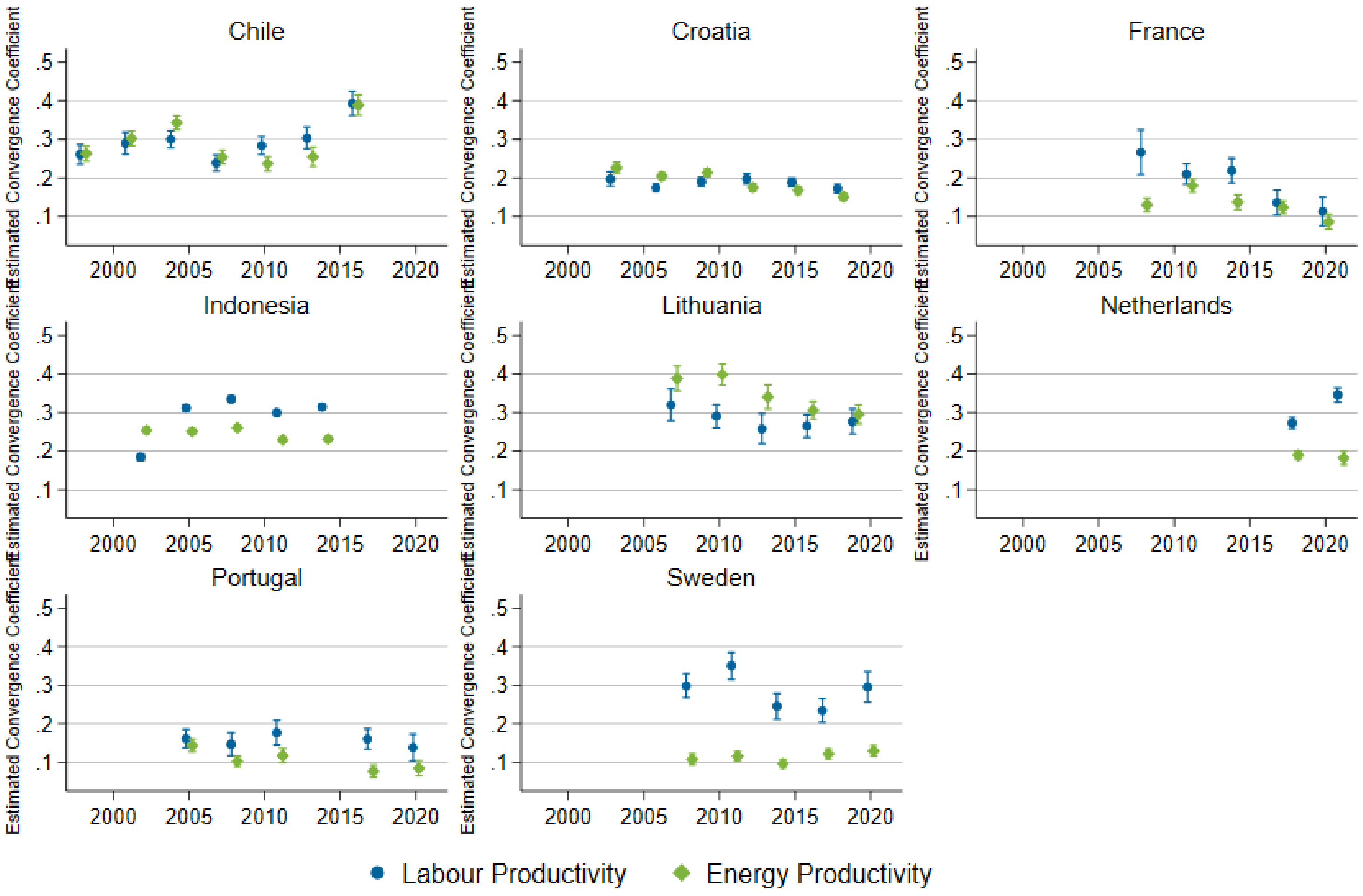
Regress firms' growth rate against distance from the frontier:

$$\Delta \log y_{ict} = \sum_{s} \beta_c^s GAP_{i,c,k(i),t-1} D_t^s + \boldsymbol{\gamma X_{ict}} + \alpha_{c,k(i),t} + \varepsilon_{ict}$$

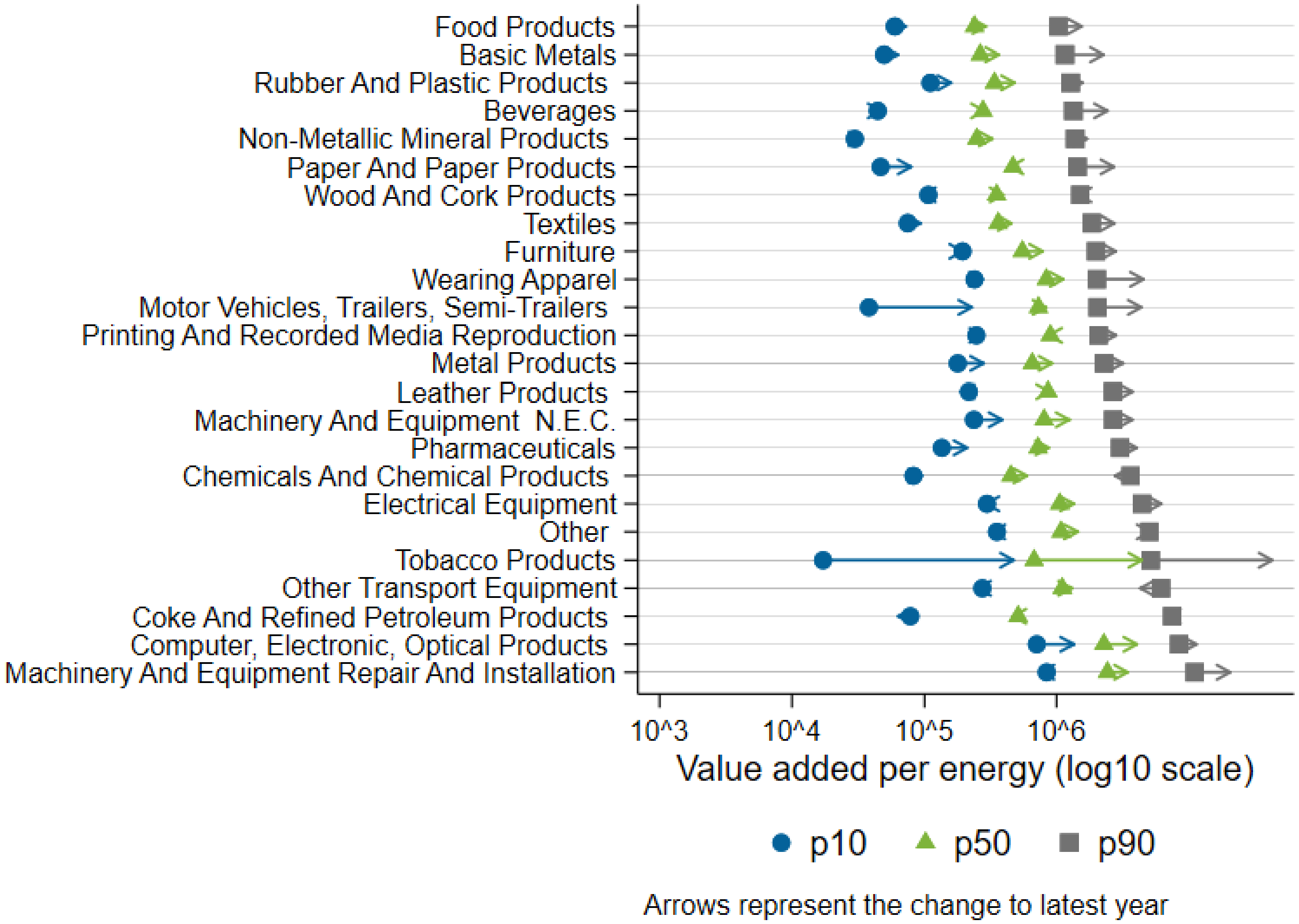
$$GAP_{i,c,k(i),t-1} = \log y_{i,c,k(i),t-1}^{frontier} - \log y_{,ic,k(i),t-1}$$

$\Delta \log y_{ict}$ is growth rate of energy or labour productivity

D_t^s is a dummy for three year periods



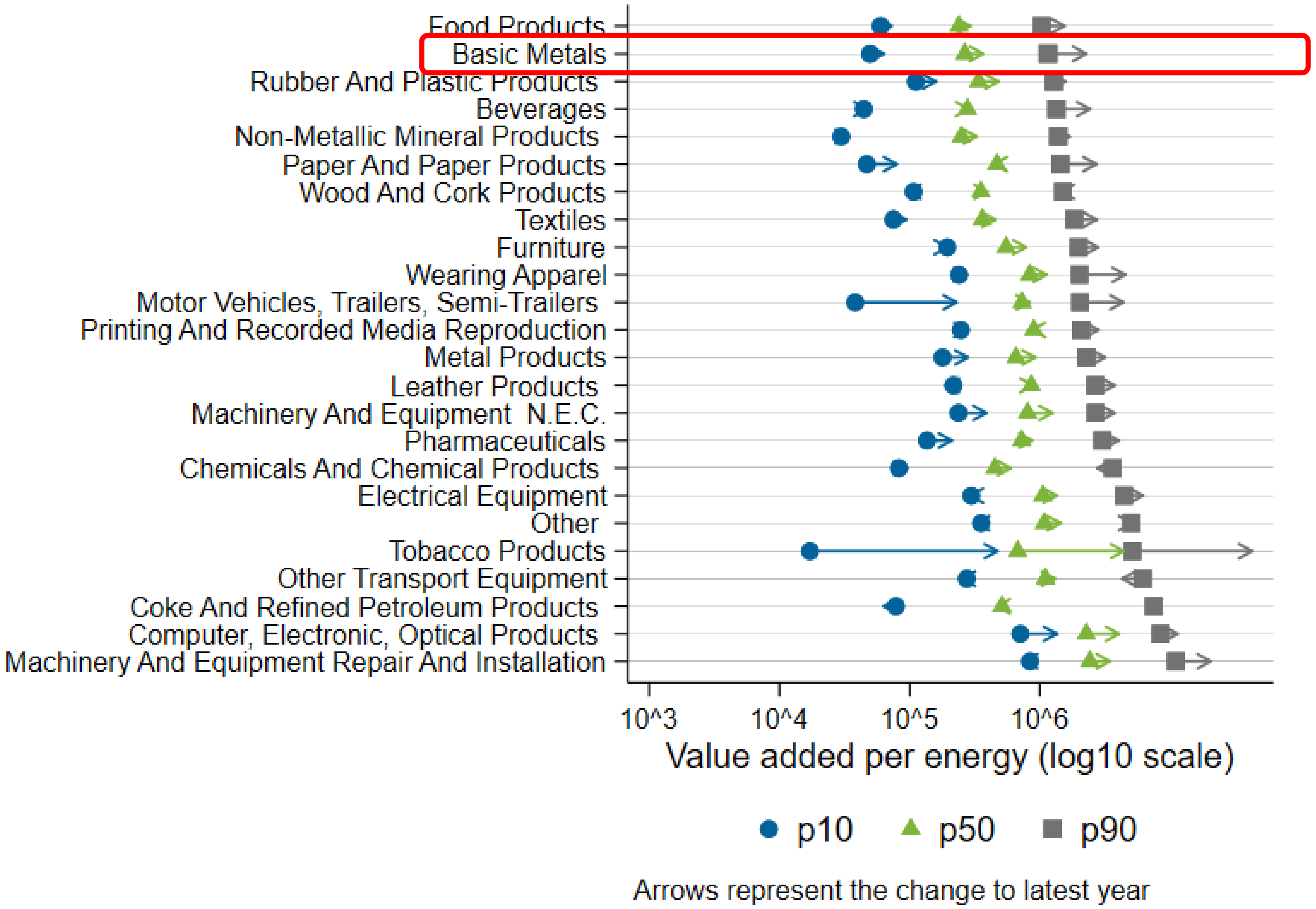
Dispersion across industries over time



Energy productivity at the 10th, 50th, and 90th percentiles within 2-digit industries and their change over time on a log scale

Computed for 3 digit industries and averaged across countries

Dispersion has increased in basic metals



Basic metals includes steel

- Homogenous output
- Heterogeneous production techniques (blast furnace and electric arc)

Energy productivity at the 10th, 50th, and 90th percentiles within 2-digit industries and their change over time on a log scale

Computed for 3 digit industries and averaged across countries

2. FACTORS EXPLAINING FIRMS' ENERGY PRODUCTIVITY

What explains differences between firms' energy productivity?

Machine learning prediction

- Insights on most important factors (e.g. like a variance decomposition)
- As used by e.g. Kleinberg et al. (2018 QJE); Bazzi et al. (2022 REStat)
- Use pooled firm-level data for Chile, Croatia, Indonesia, Lithuania, Portugal
- XGBoost: extreme gradient boosting
  - Combines many weak prediction methods like decision trees
- Train model on 80% of data, test on 20%
- Subtract country-industry-year fixed effects within the test/train sample
- SHAP analysis on random sample of 5 000 observations to analyse model features

Linear regression

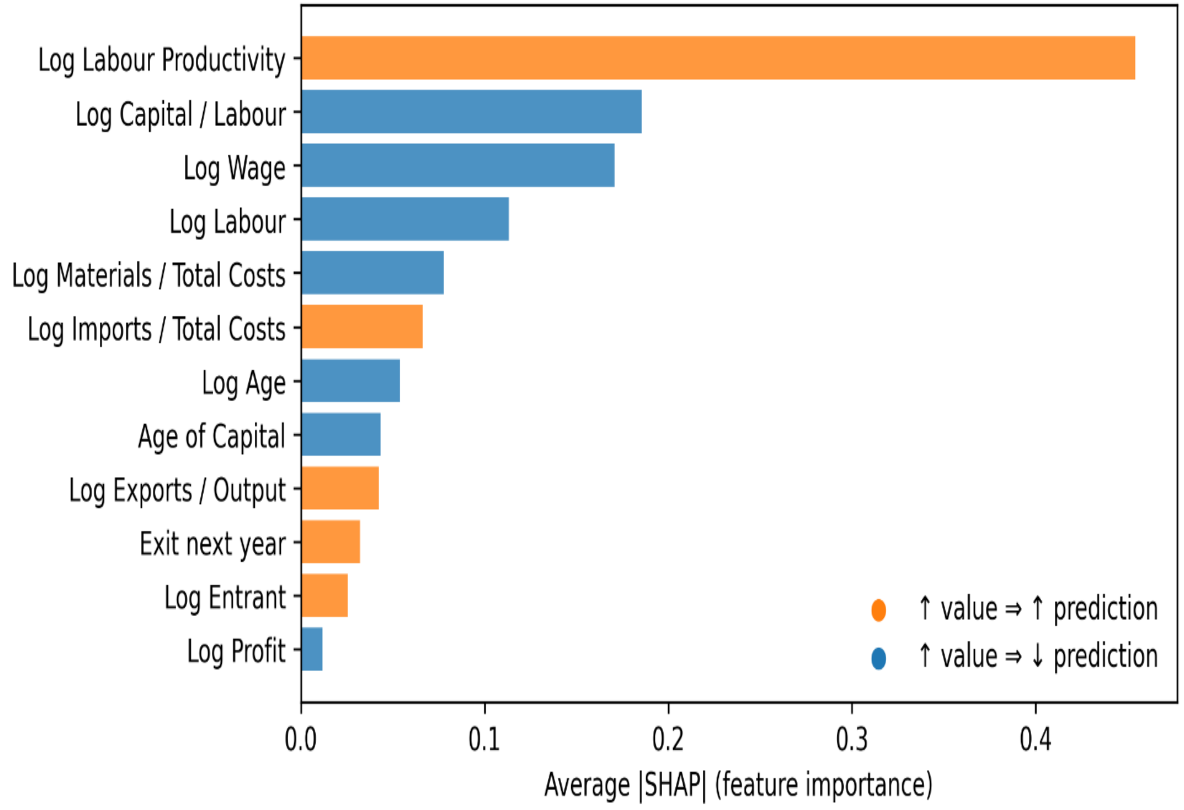
- Insights on direction and magnitude of correlations
- Estimated separately for each country
- Estimate:

$$\log\left(\frac{VA_{ict}}{ENQ_{ict}}\right) = \beta_c\, x_{\text{ict}} + \alpha_{ck(i)t} + \varepsilon_{ict}$$

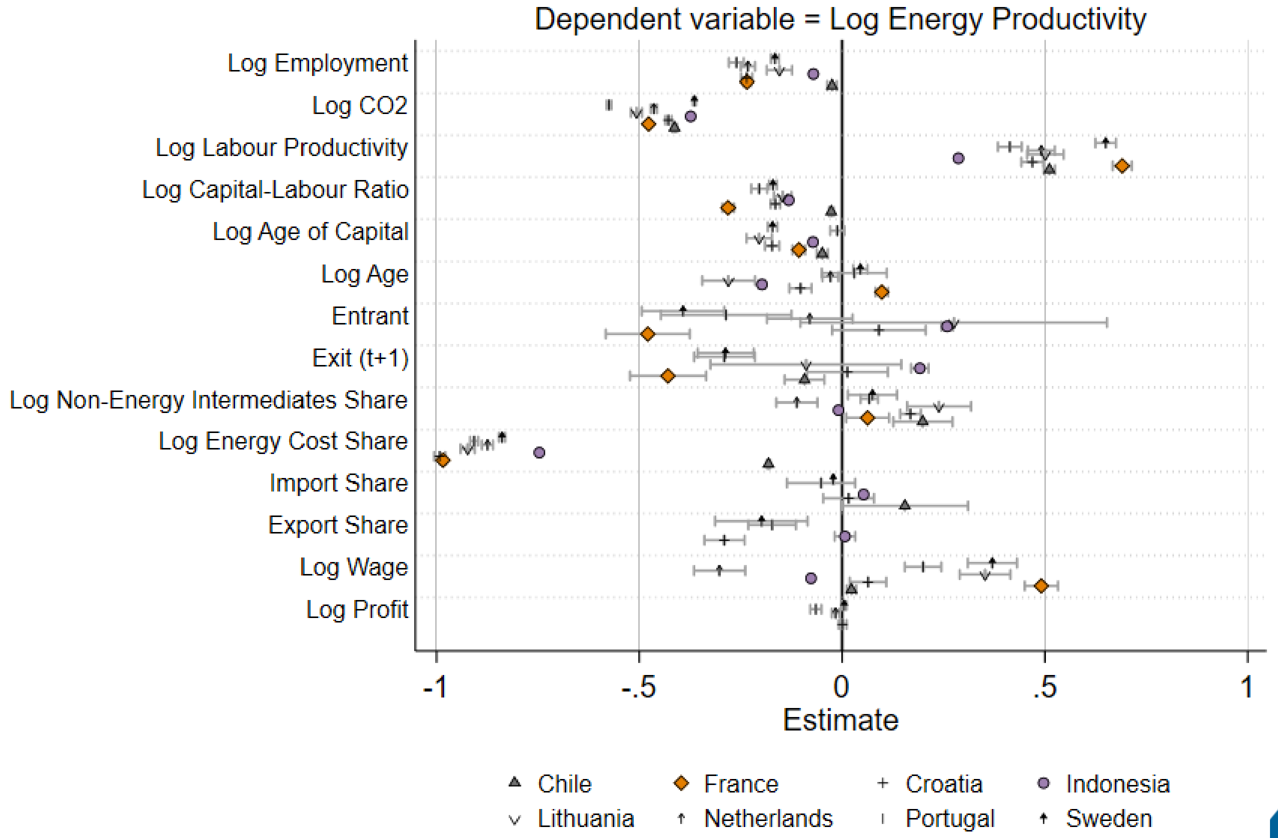
- x_{ict} is each explanatory variable
- $\alpha_{ck(i)t}$ is a 4 digit industry-year fixed effect
- Baseline: bivariate regressions with each x_{ict} separately
- Robustness: multivariate with all x_{ict} together

Results from ML gradient boosting and OLS

ML Prediction: Feature importance

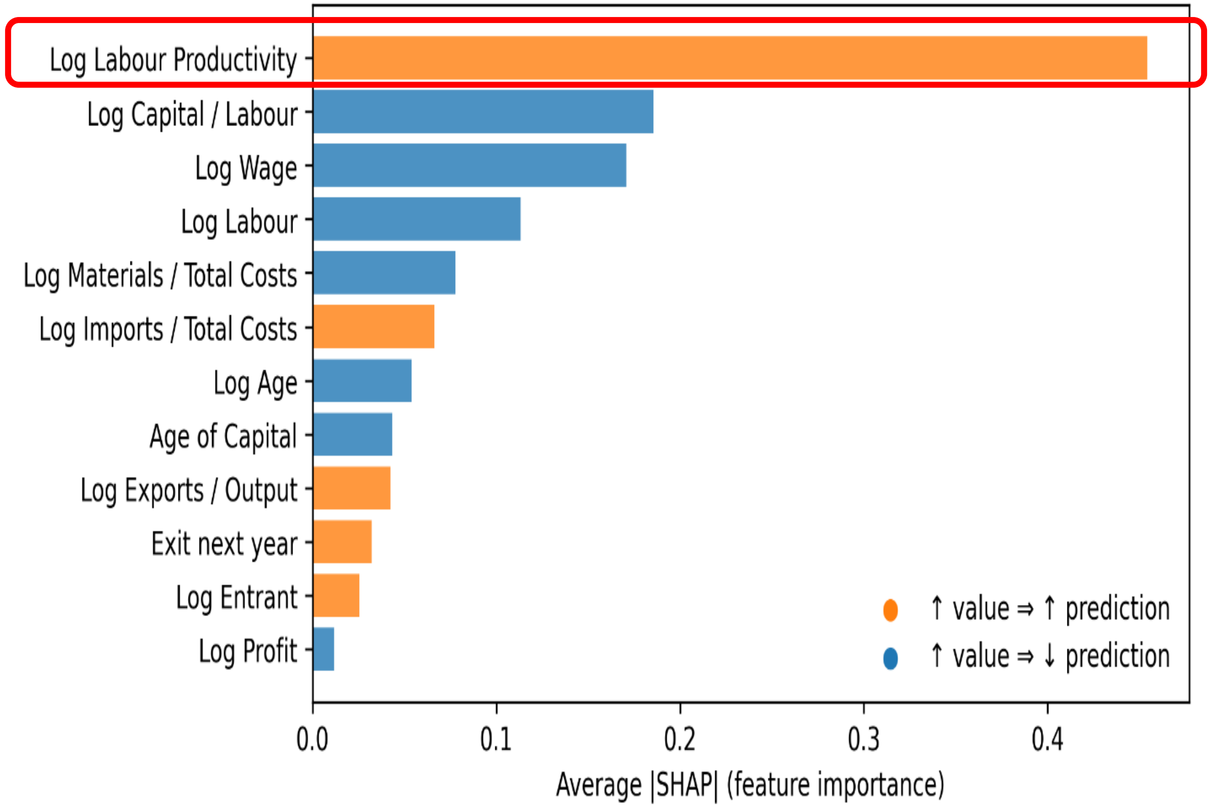


Bivariate OLS: estimated coefficients

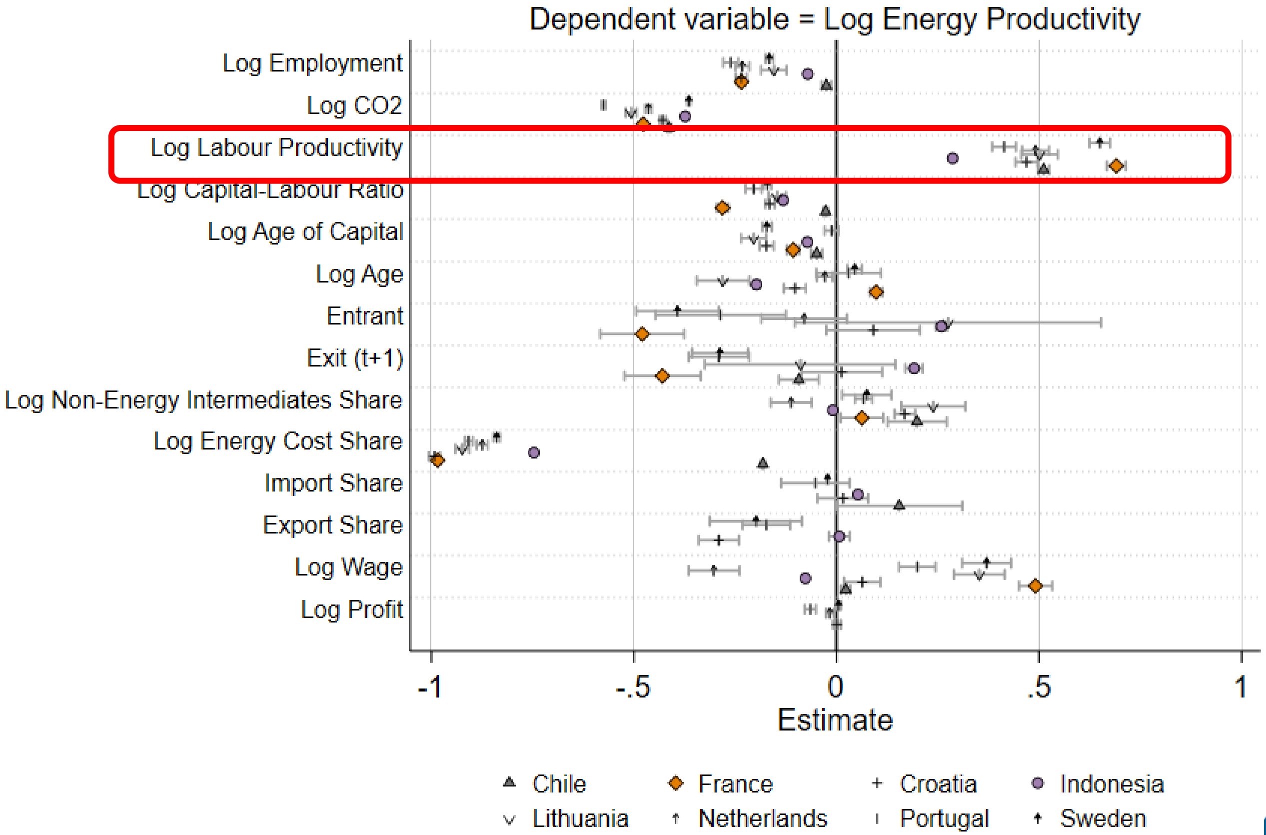


Labour productivity is the most important predictor of energy productivity

ML Prediction: Feature importance



Bivariate OLS: estimated coefficients

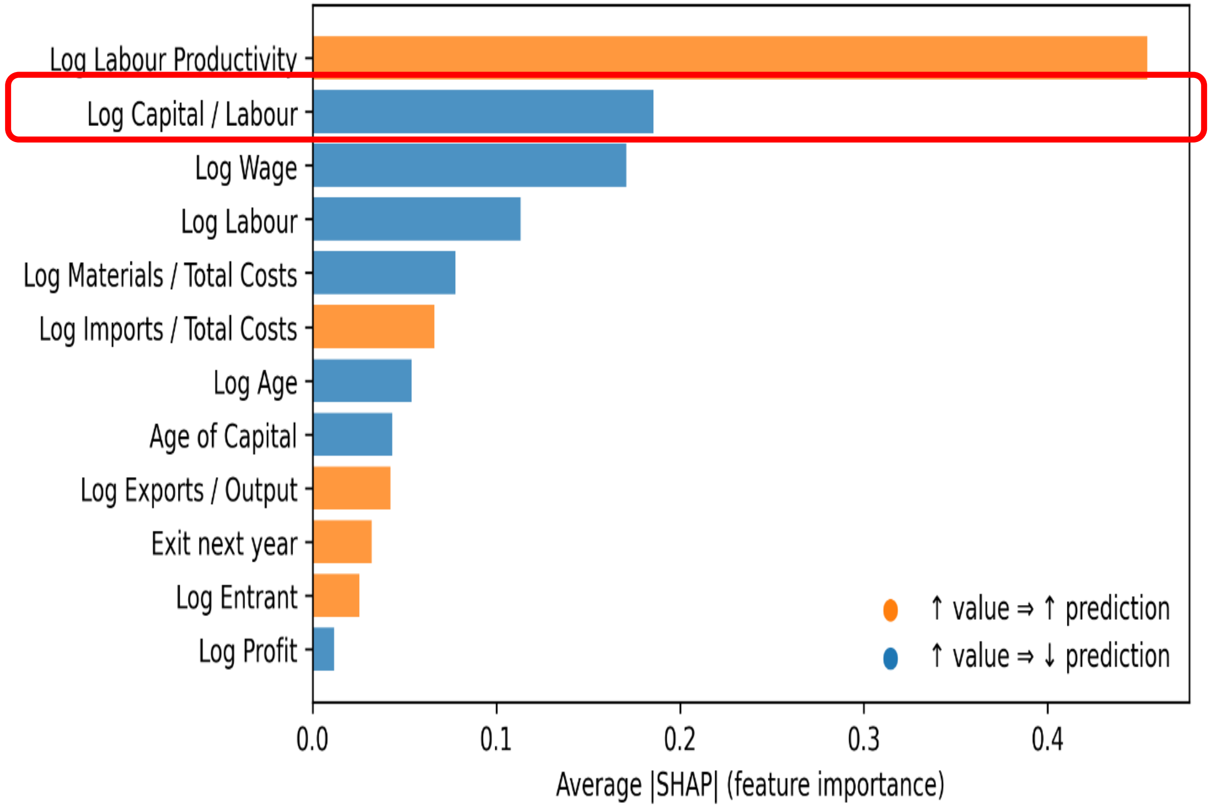


- This holds controlling for all other variables (also within-firm changes)
- For ML prediction, with labour productivity as the target, wage is the most important variable

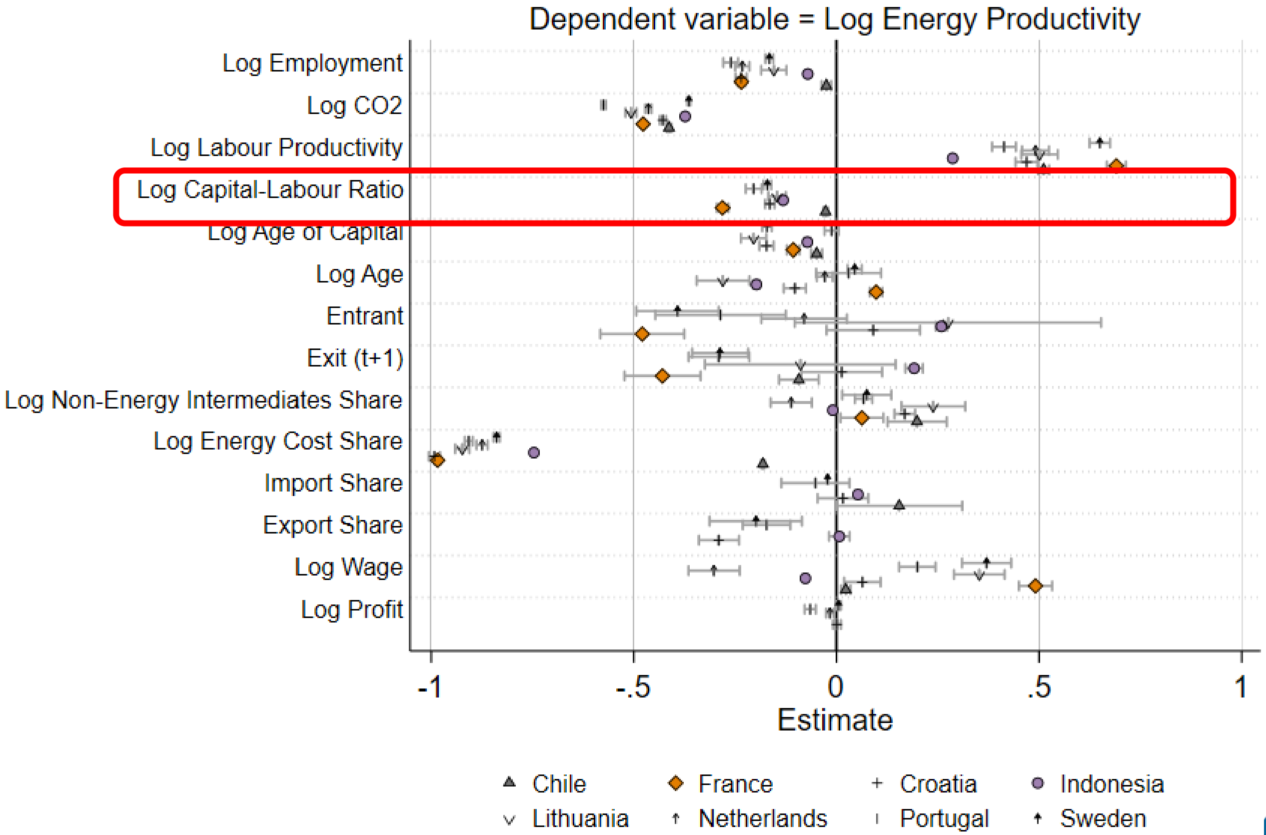
➢ **Suggests a positive relationship between labour and energy productivity**

Capital intensity is negatively associated with energy productivity

ML Prediction: Feature importance



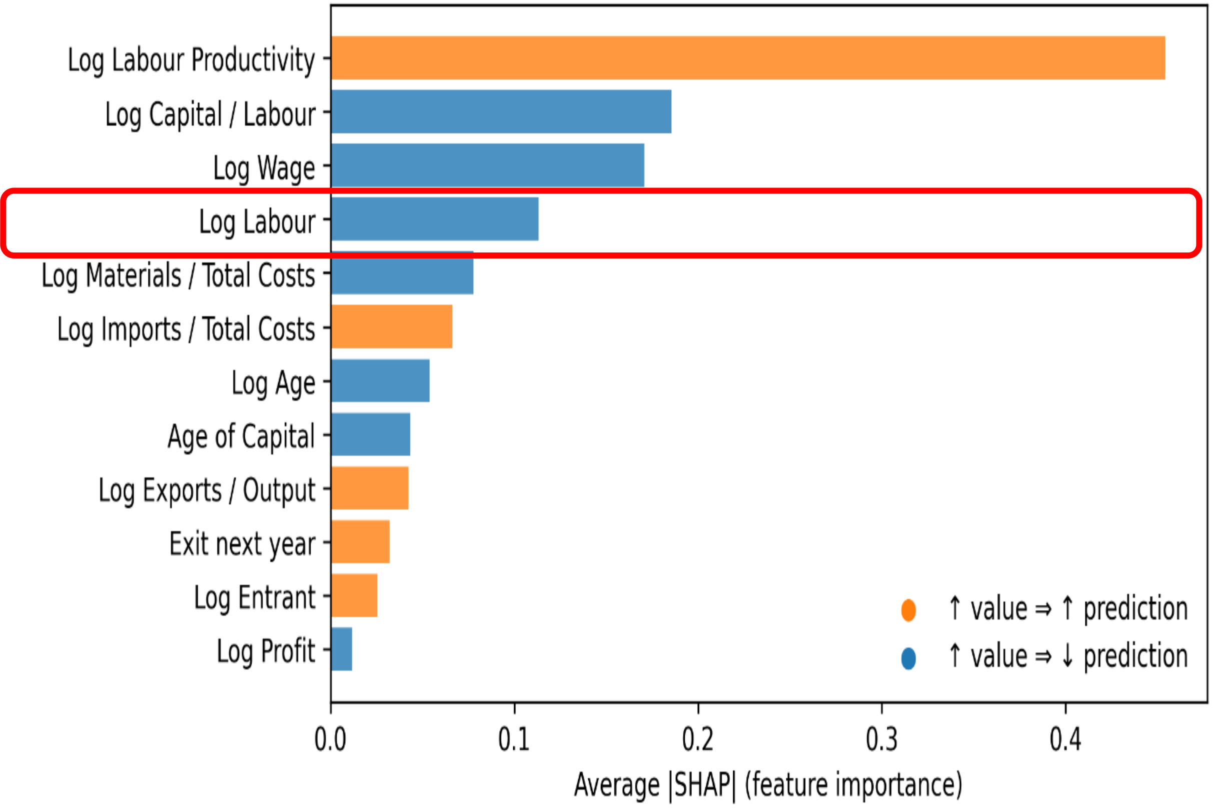
Bivariate OLS: estimated coefficients



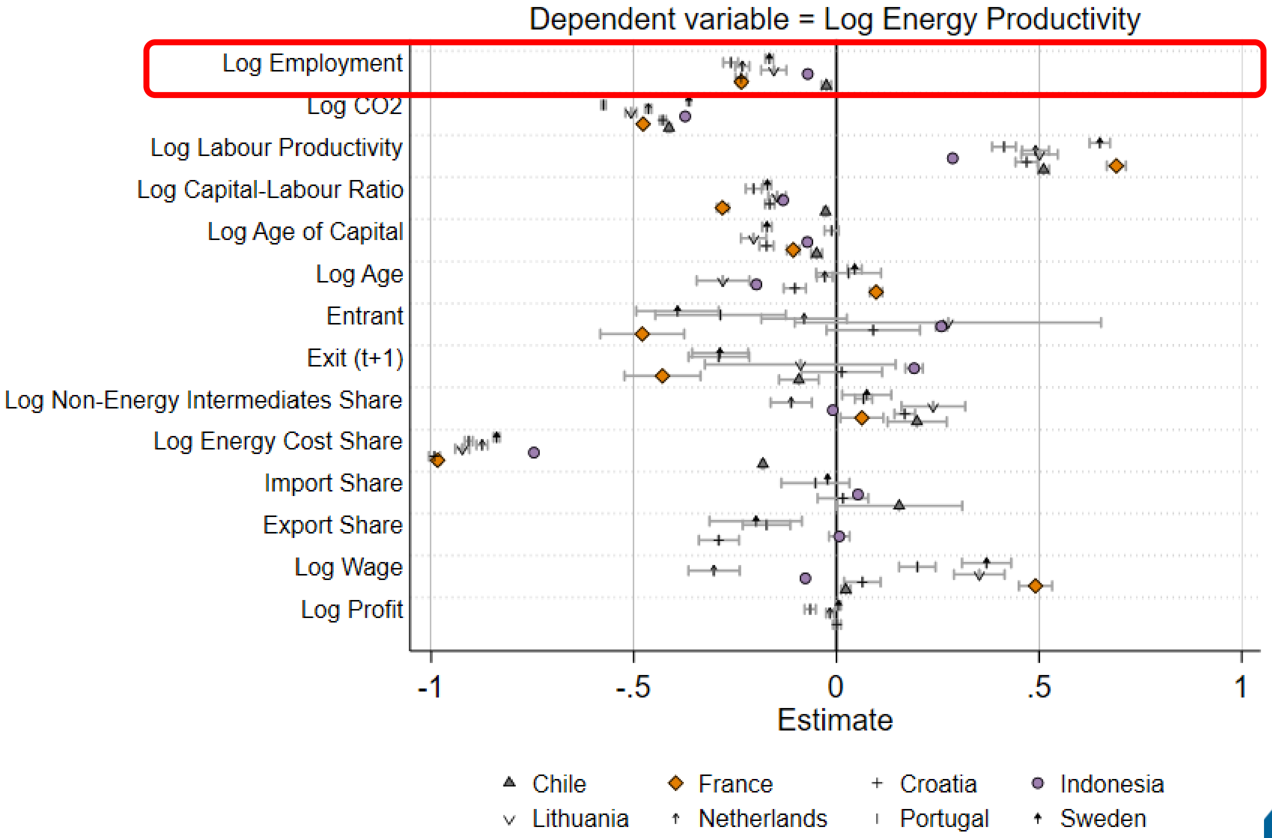
- Holds in multivariate setting
- ➢ **Evidence of capital-energy complementarity**

Larger firms are less energy productive

ML Prediction: Feature importance



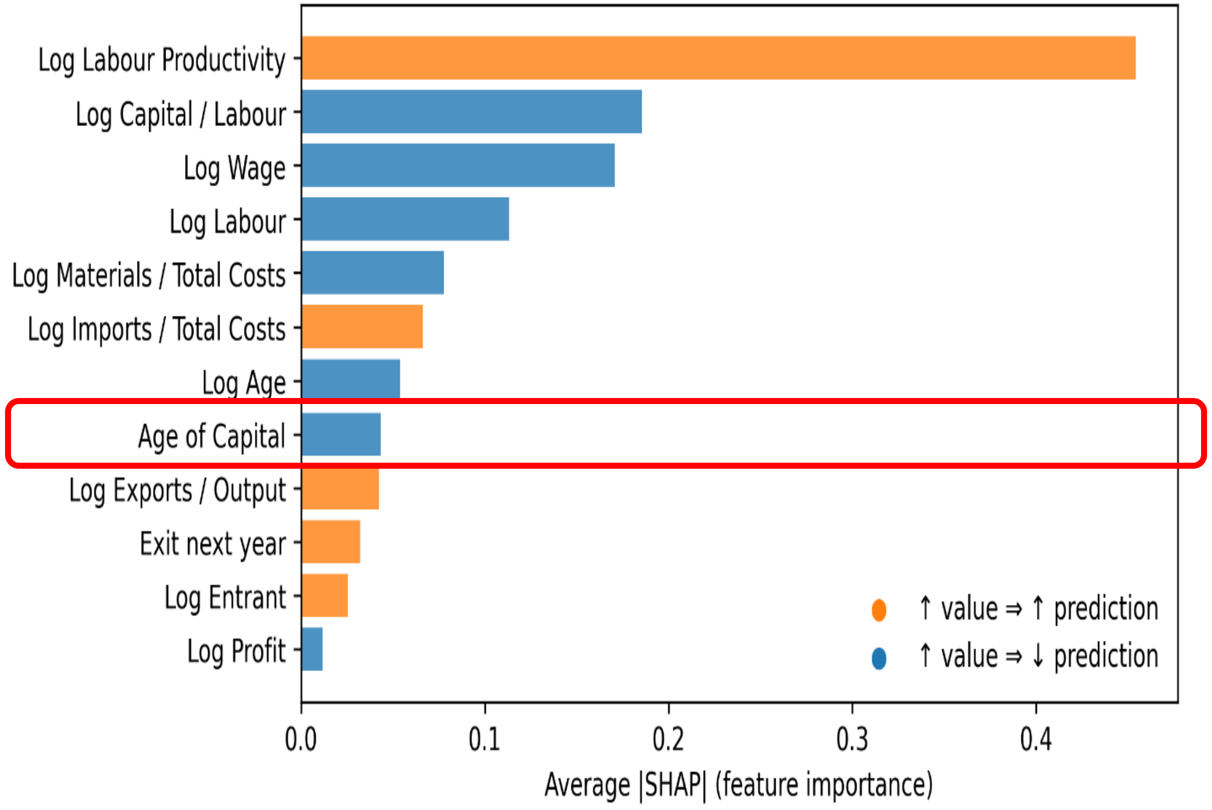
Bivariate OLS: estimated coefficients



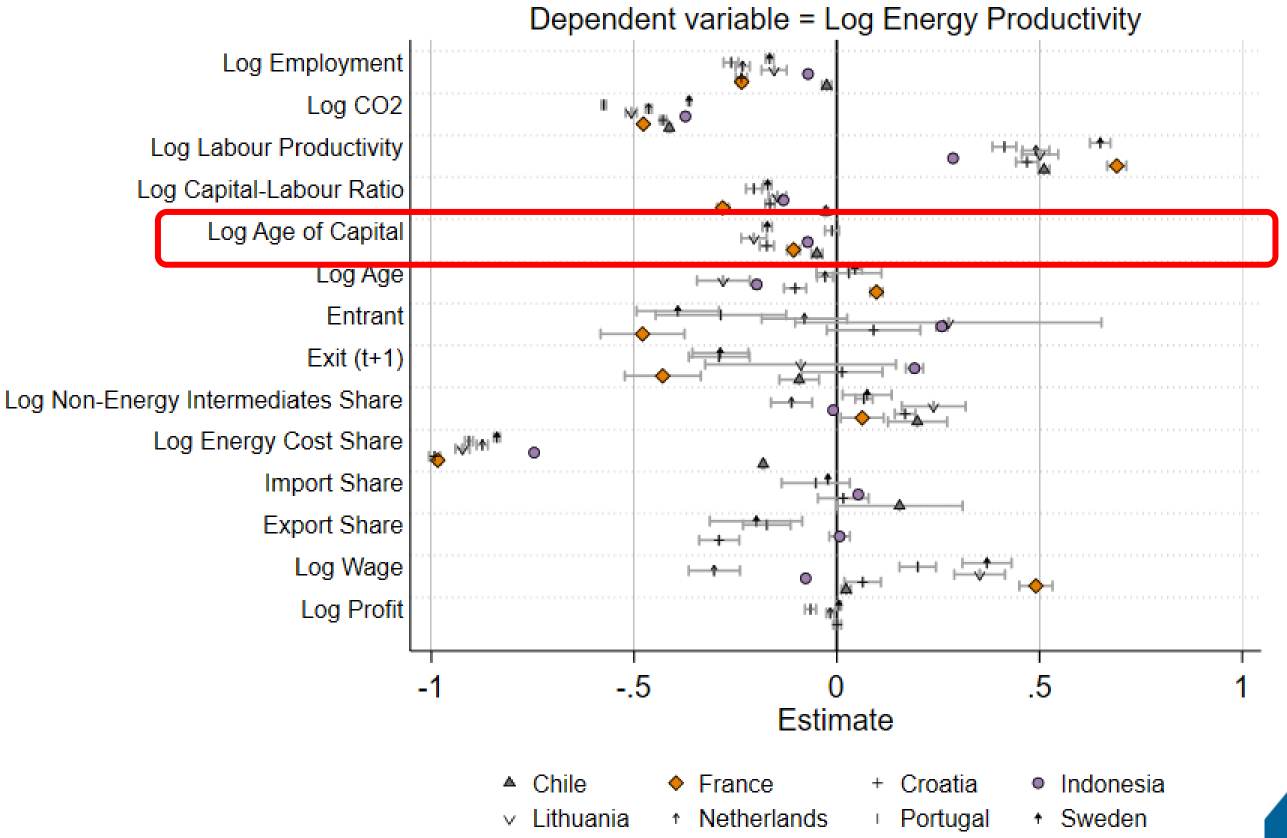
- Holds controlling for other factors
- ➢ **In stark contrast to strong positive relationship between size and labour productivity or TFP**
- ➢ Possible explanations? **Significantly lower energy prices for large firms**; lack of scale economies for energy

Energy productive firms use younger capital

ML Prediction: Feature importance



Bivariate OLS: estimated coefficients

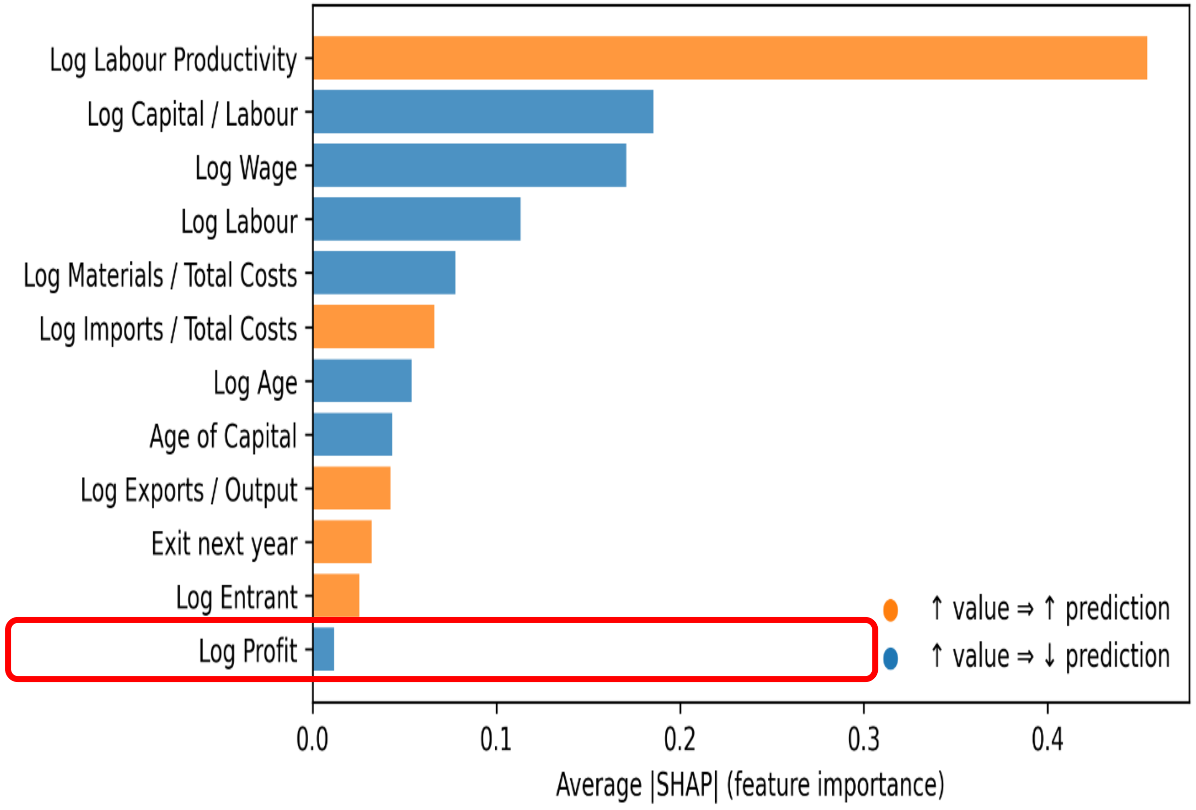


- Age of capital defined as years since investment of 20% capital stock, following Fiori and Scoccianti (2025)
- Holds with controls, including for overall capital intensity

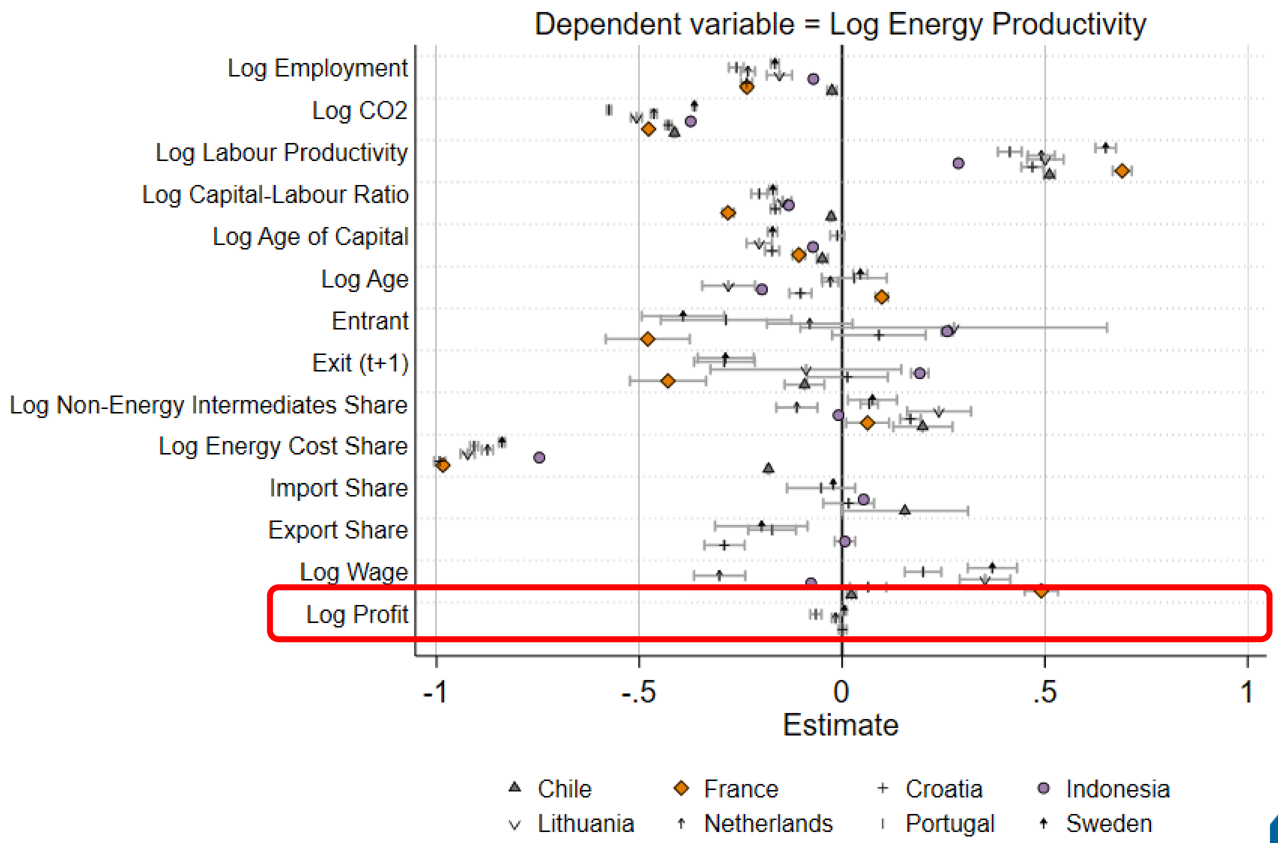
➢ **New vintages of capital appear to be more energy efficient**

Profit has almost no power in predicting energy efficiency

ML Prediction: Feature importance



Bivariate OLS: estimated coefficients



- Much more important for predicting labour productivity

➢ **Average share of energy costs in total costs = 5%; labour costs = 31%**

3. THE AGGREGATE IMPLICATIONS OF DISPERSION

Improving energy productivity of laggards would yield massive energy and emissions reductions

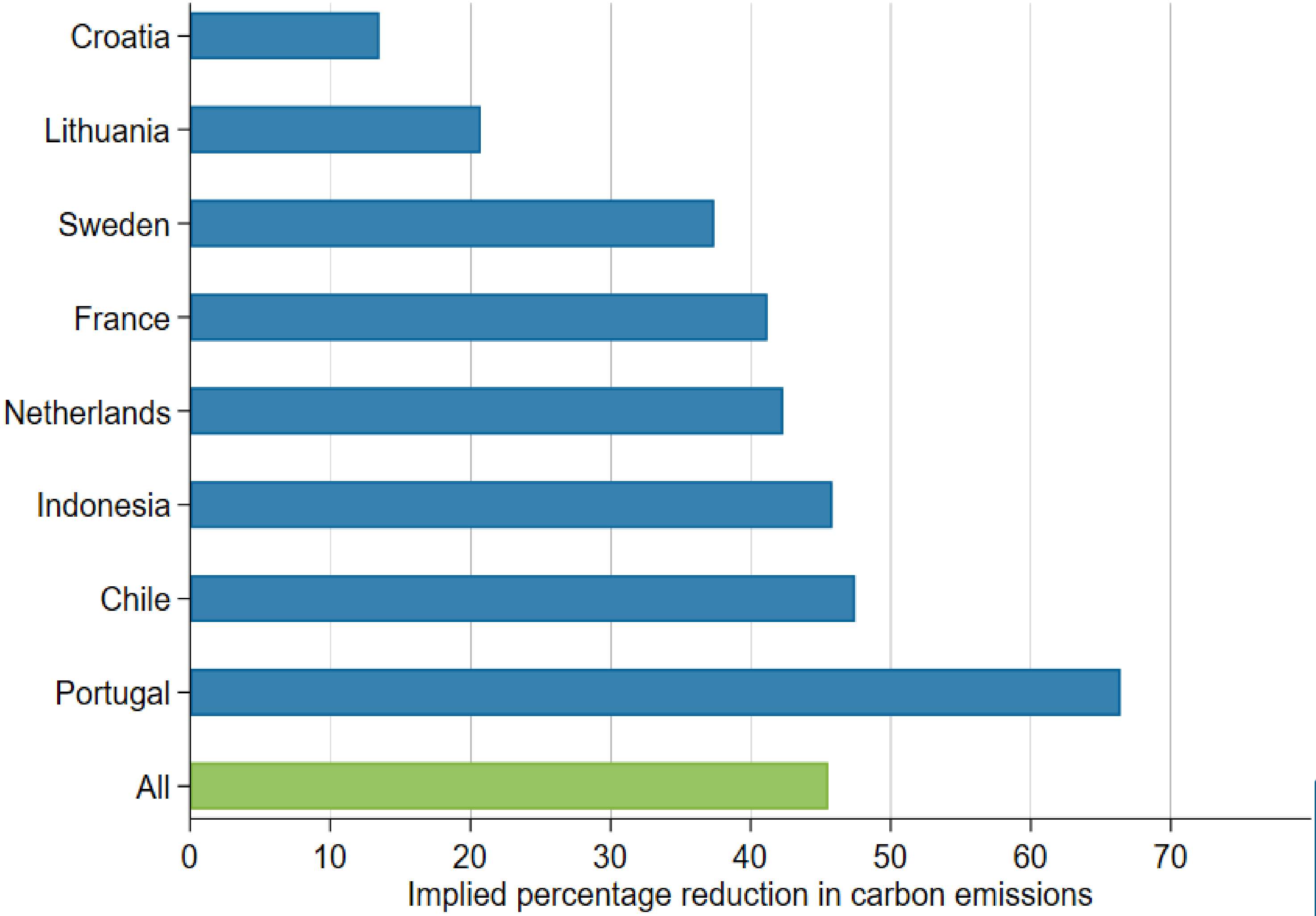
<u>Counterfactual exercise</u>

- Improve energy productivity of unproductive firms to the 25th percentile in their industry
- Compute aggregate energy needed to achieve same value added

➢ **45% lower energy for same value added**

  – Equivalent for labour productivity = 8%

- Why so large?
  – Huge dispersion – tail is far behind
  – Unproductive firms are larger



As industries become more dispersed in energy productivity, they tend to become less energy productive

- $\Delta_5 \log(y_{ckt}) = \beta\ \Delta_5 DispEP_{ckt} + \varepsilon_{ckt}$
- Δ_5 denotes 5 year changes
- Where: $DispEP_{ckt} = \log\left(\frac{VA_{ckt}}{ENQ_{ckt}}\right)^{p90} - \log\left(\frac{VA_{ckt}}{ENQ_{ckt}}\right)^{p10}$
- Country c, industry k, year t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(VA)	log(Energy)	log(CO_2)	log(L)	Log(VA/Energy)	log(VA/CO_2)	log(VA/L)
$\Delta_5 DispENQ$	-0.0306*	0.0657***	0.0507***	-0.0116	-0.0966***	-0.0795***	-0.0194
	(0.0181)	(0.0179)	(0.0189)	(0.00993)	(0.0163)	(0.0174)	(0.0140)
Observations	3195	3147	3125	3200	3134	3104	3195

4. WHAT EXPLAINS DISPERSION IN ENERGY PRODUCTIVITY?

What explains dispersion in energy productivity?

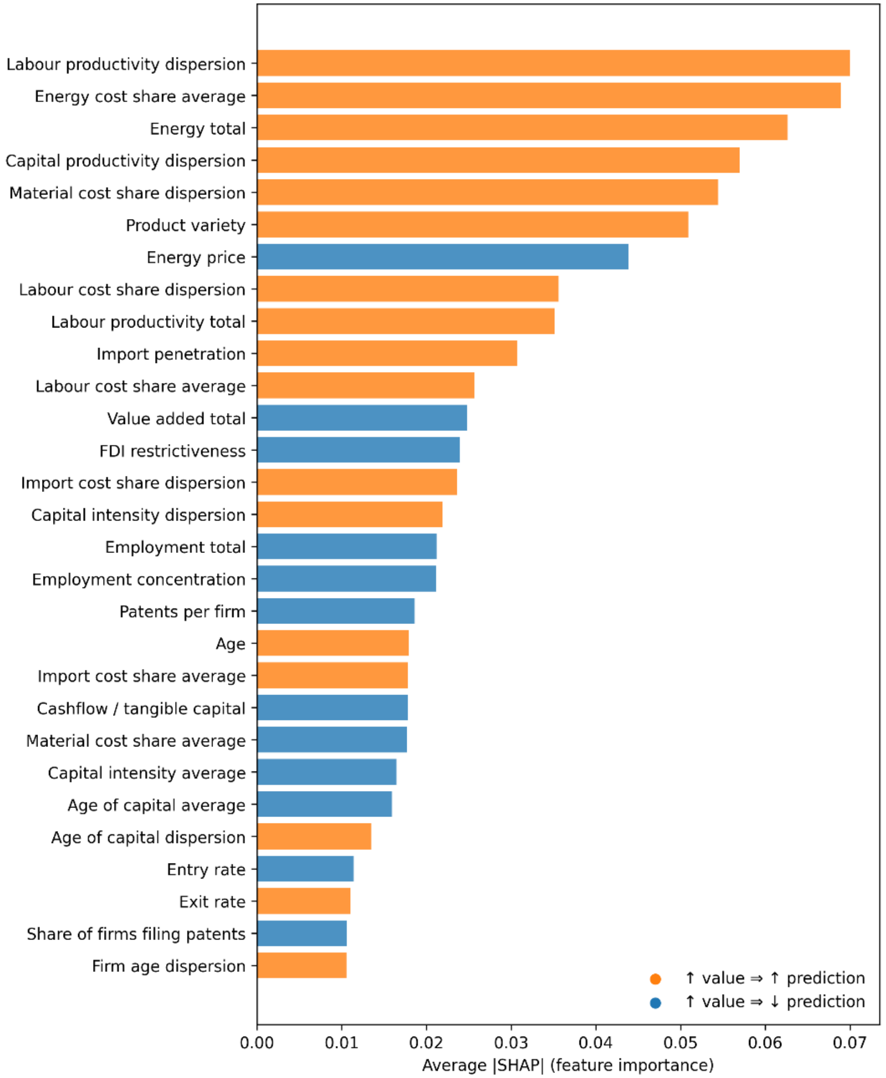
- Use the same ML prediction model method as before, alongside OLS
- For country-industry-year data, missing values on some variables
  - E.g. due to confidentiality
- ML has the advantage of including observations with missing values
  - Improves performance relative to multivariate OLS (in terms of RMSE and R2)
- For baseline OLS, estimate bivariate regressions:

$$DispEP_{ckt} = \beta\, x_{ckt} + \alpha_{ct} + \varepsilon_{ckt}$$

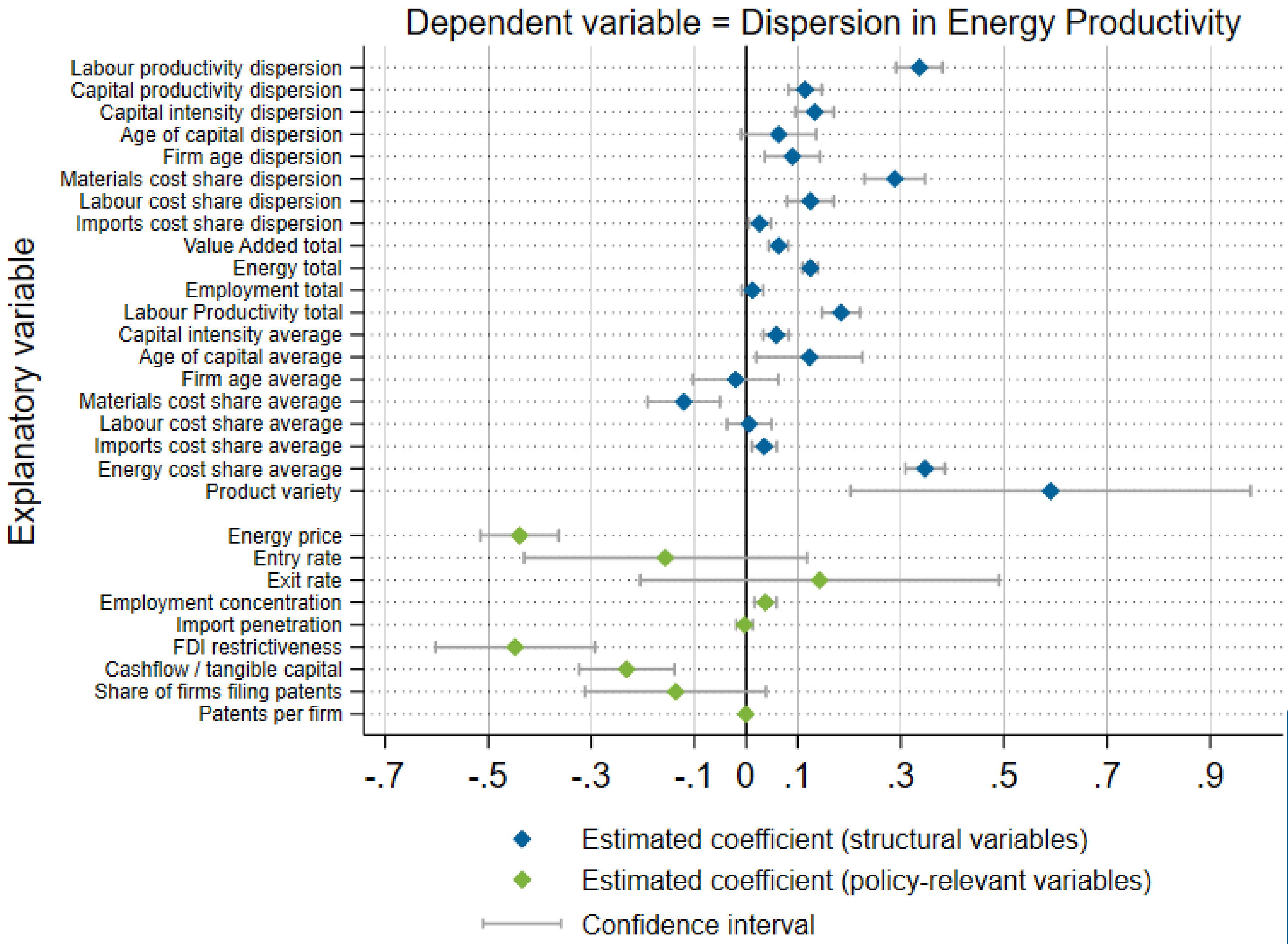
- Robustness with multivariate
- Also estimated within-unit differences: $\Delta_5 DispEP_{ckt} = \beta\, \Delta_5 x_{ckt} + \alpha_{ct} + \varepsilon_{ckt}$

Factors explaining dispersion in energy productivity

ML Prediction: Feature importance

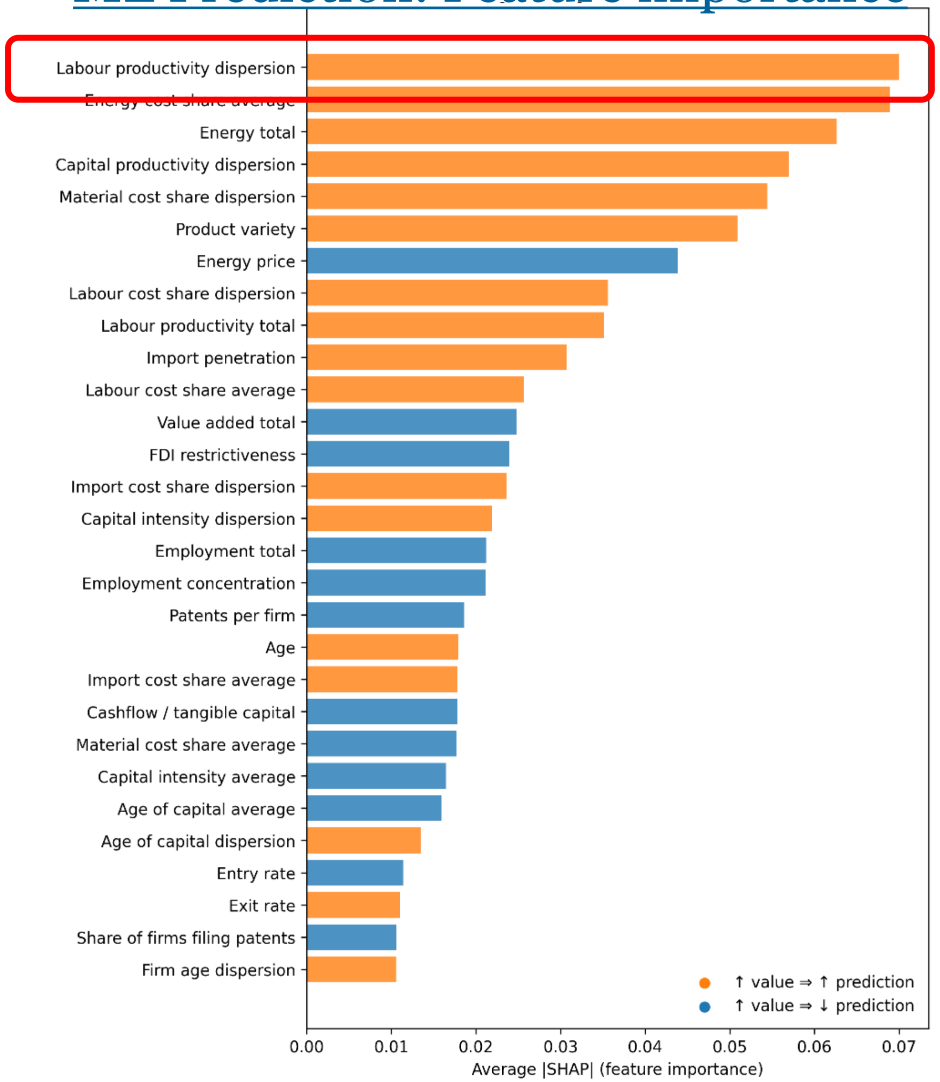


Bivariate OLS: estimated coefficients

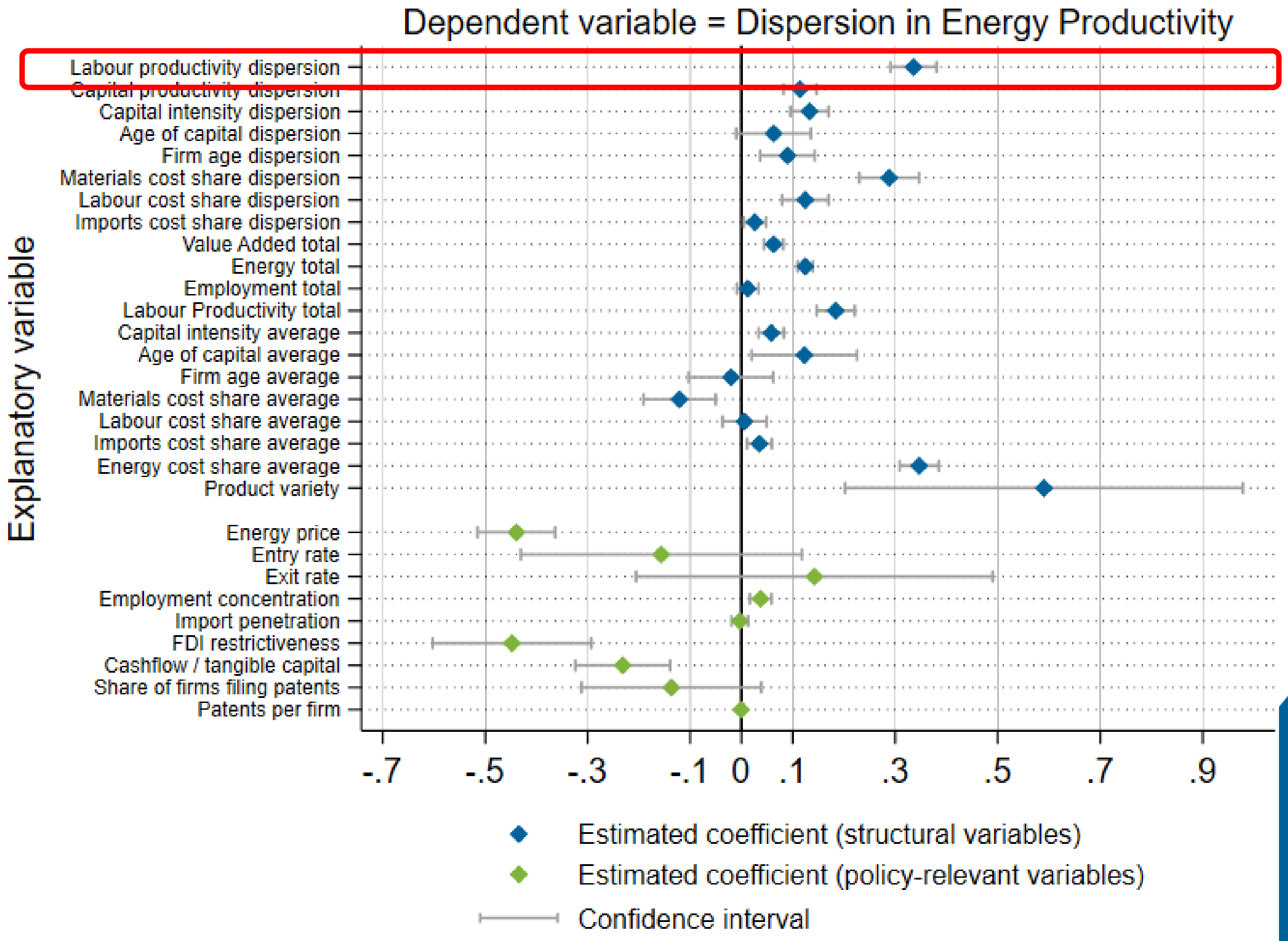


Dispersion in labour and energy productivity are strongly related

ML Prediction: Feature importance

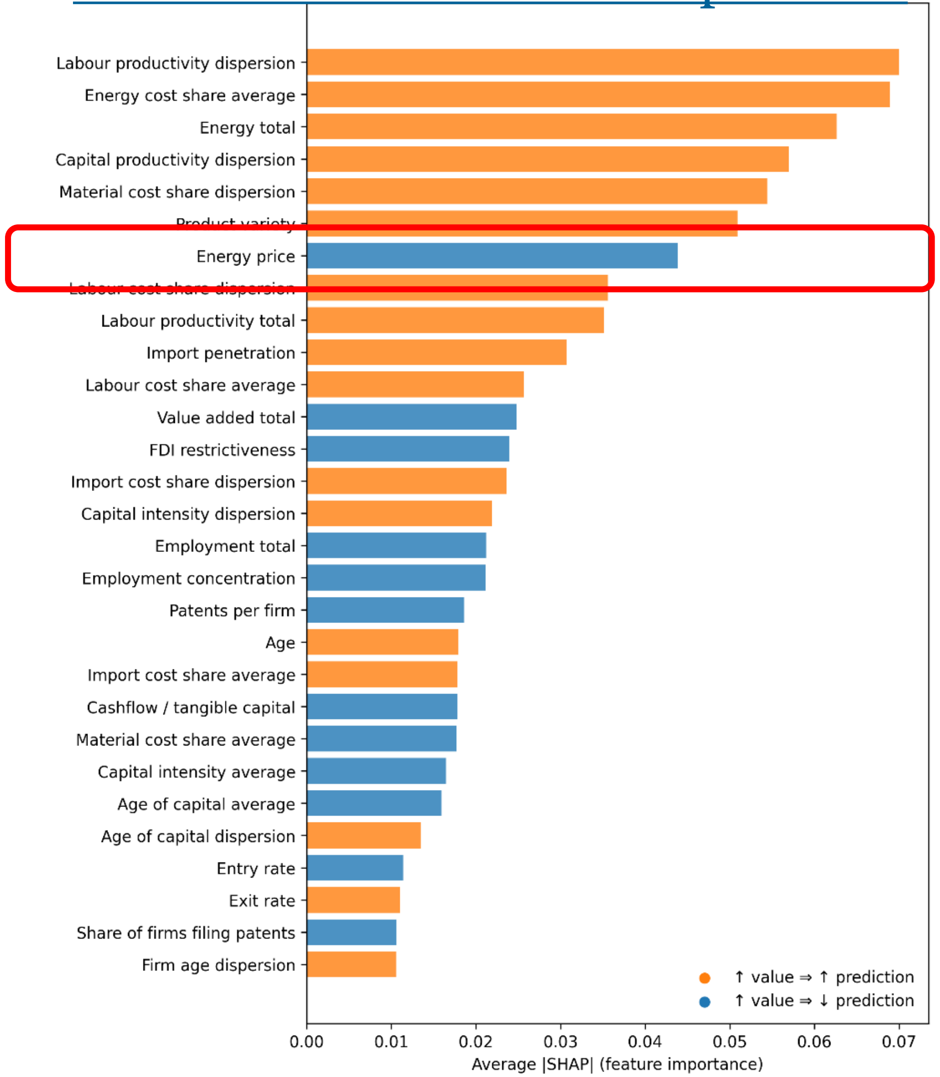


Bivariate OLS: estimated coefficients

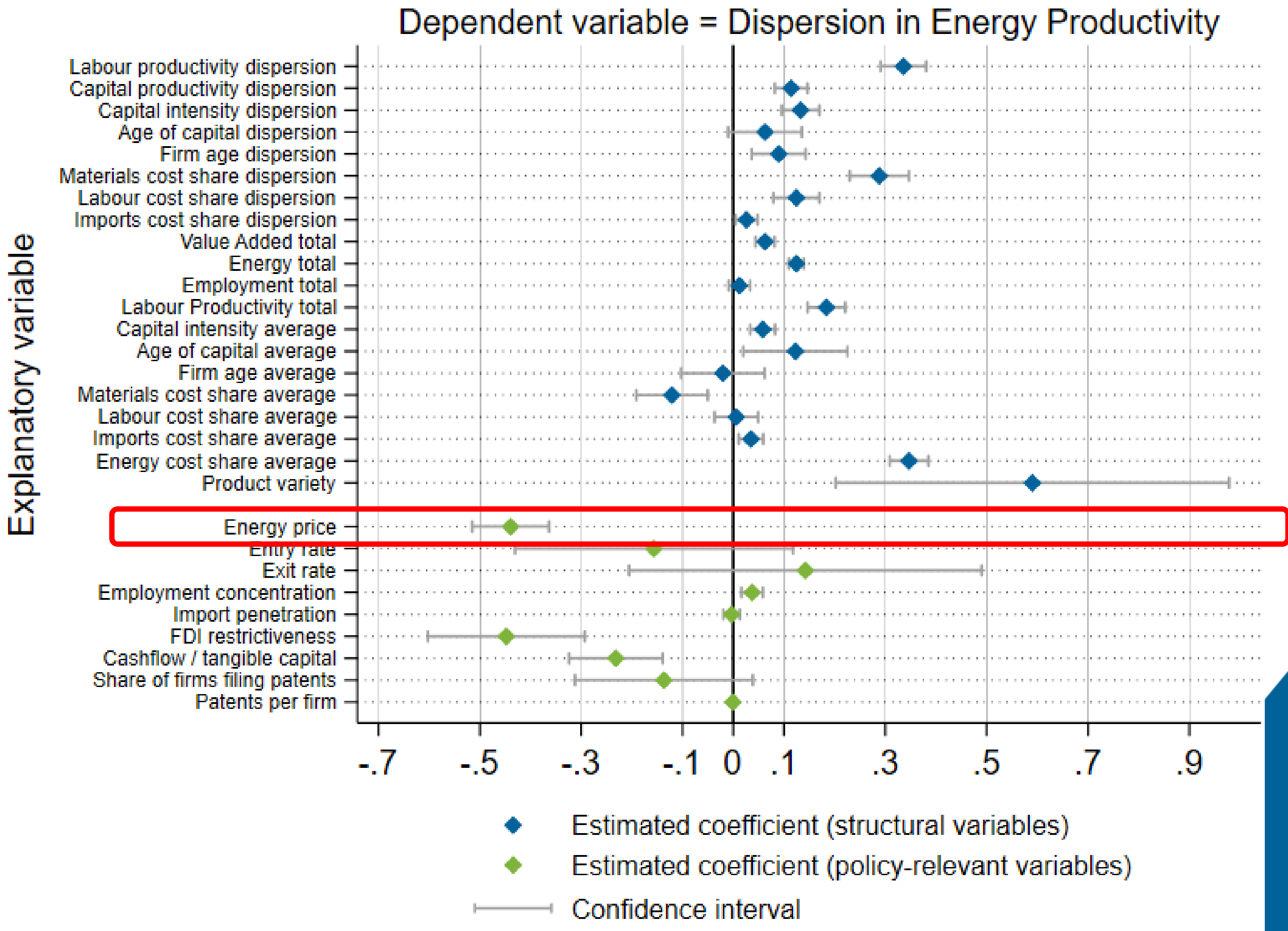


Higher energy prices associated with lower and declining dispersion

ML Prediction: Feature importance

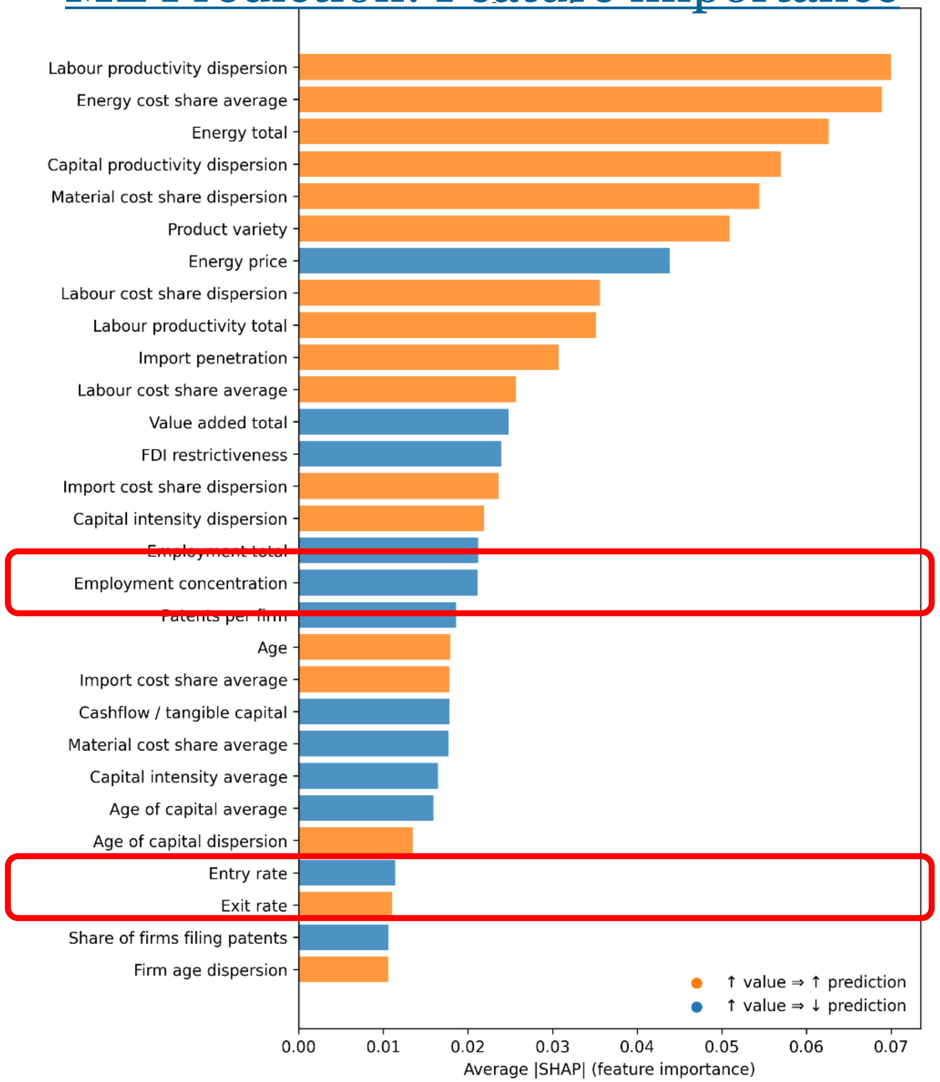


Bivariate OLS: estimated coefficients

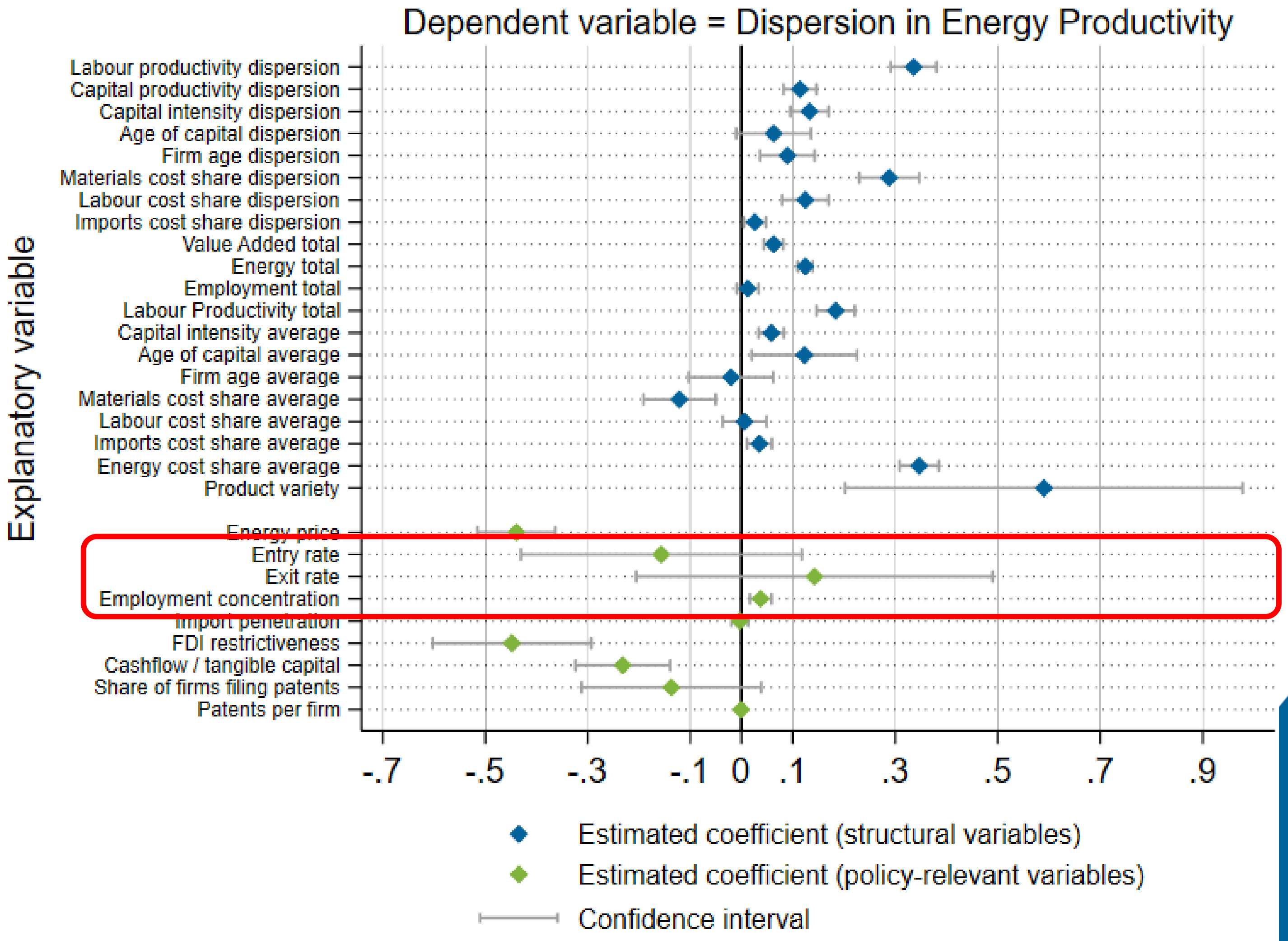


Competition and business dynamism may reduce dispersion

ML Prediction: Feature importance

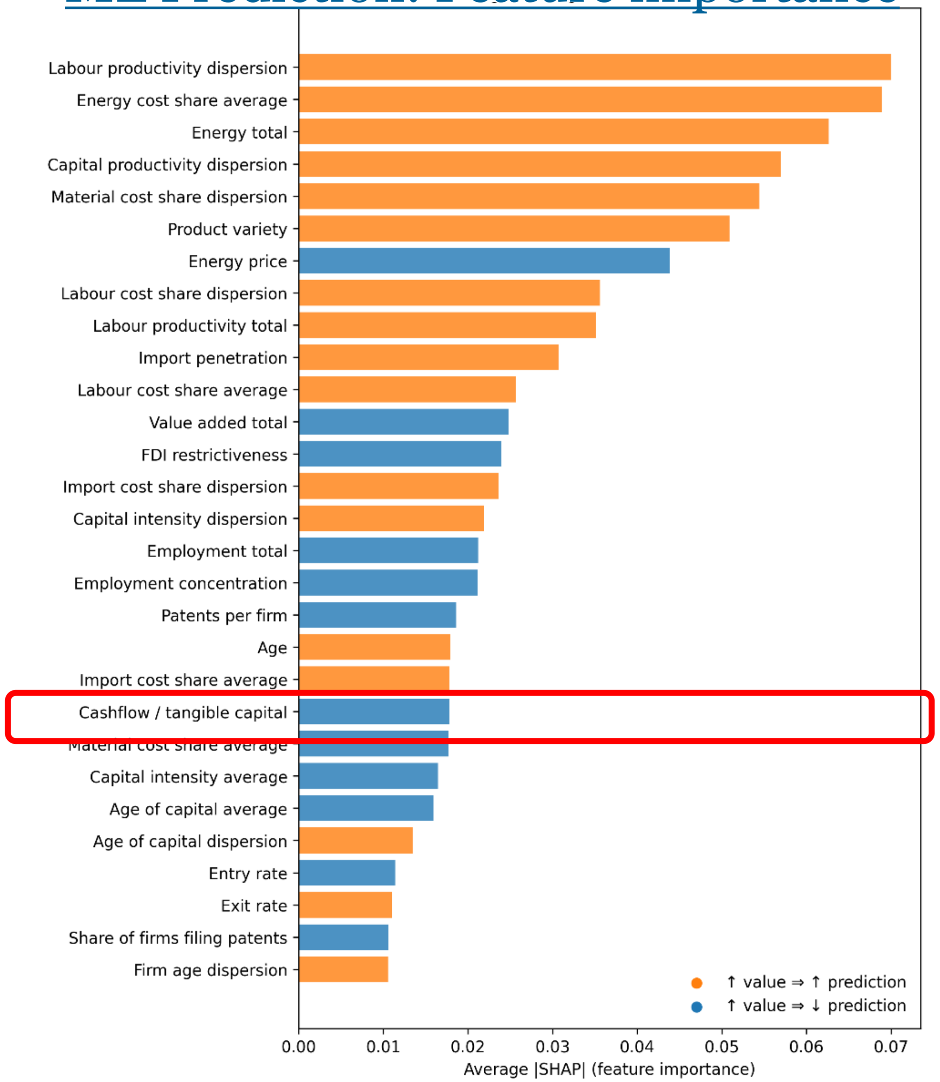


Bivariate OLS: estimated coefficients

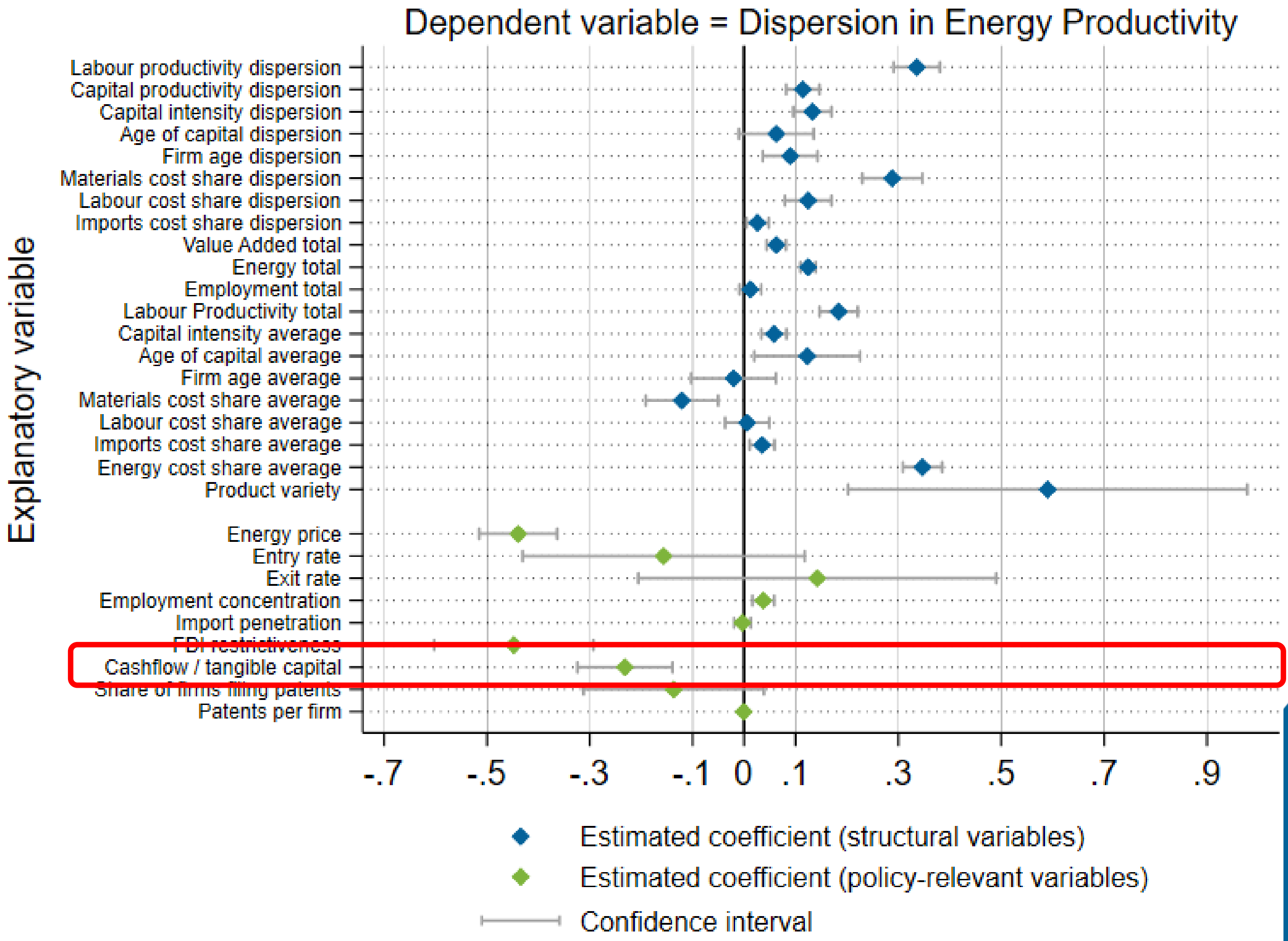


Greater access to finance associated with less dispersion

ML Prediction: Feature importance

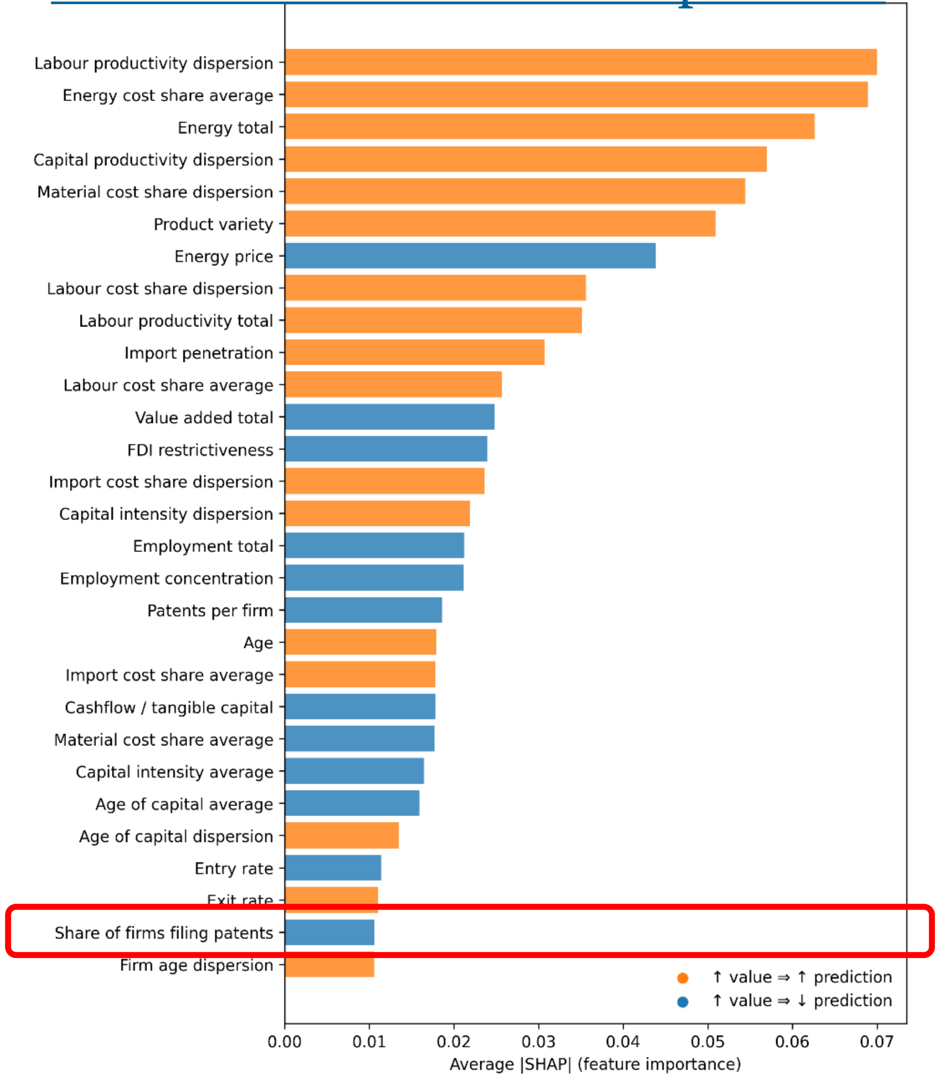


Bivariate OLS: estimated coefficients

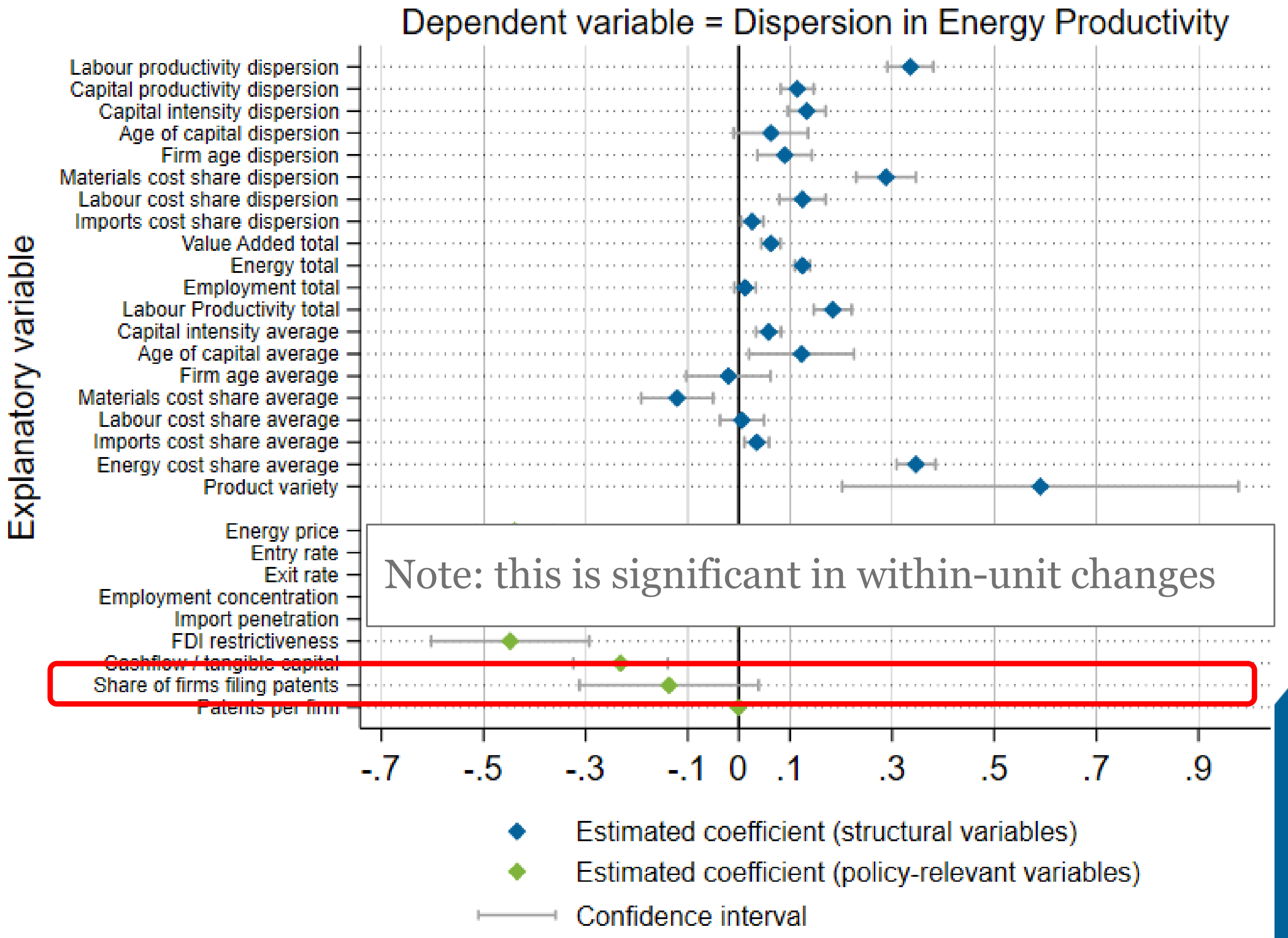


More firms innovating associated with declining dispersion

ML Prediction: Feature importance



Bivariate OLS: estimated coefficients



CONCLUSIONS

Conclusions and policy implications

Key findings

1. Huge dispersion in energy productivity between firms within industries
2. Economic and energy productivity strongly linked at firm level
3. Improving productivity of laggards has great potential to reduce energy and emissions
4. Energy prices, competition, and access to finance associated with lower dispersion

Policy implications

1. Policies targeting least energy efficient firms key for energy savings and net zero
2. Helpful policies go well beyond energy pricing
3. Improving energy and labour productivity can be complementary

THANKS!

josh.delyon@oecd.org
antoine.dechezlepretre@oecd.org

BETTER POLICIES FOR BETTER LIVES

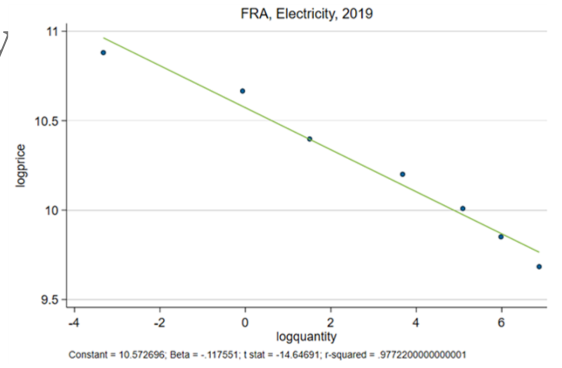
APPENDIX

Data sources by country

Country	Years covered	Type of energy use data	Energy use data sample	Production data sample	Unit level
Chile	1995-2015	Quantity by type of energy	All establishments with at least 10 employees	All establishments with at least 10 employees	Establishment, converted to firm level
Croatia	2002-2019	Total spending on energy	All firms	All firms	Firm
France	2005-2020	Quantity by type of energy	Stratified sample for establishments between 20 and 250 employees with survey weights. All establishments above 250 employees	All firms	Establishment, converted to firm level
Indonesia	2000-2019	Quantity by type of energy	All establishments with at least 20 employees	All establishments with at least 20 employees	Establishment
Lithuania	2004-2020	Total spending on energy	Sample - weights constructed using business register	All firms	Firm
Netherlands	2015-2021	Total spending on energy	Stratified sample with survey weights	Stratified sample with survey weights (same source as energy variable)	Firm
Portugal	2004-2020, exc. 2012-2014	Quantity by type of energy	Stratified sample with survey weights	All firms	Firm
Sweden	2005-2021	Quantity by type of energy	All firms with at least 10 employees	All firms with at least 10 employees	Firm

Estimating emissions with total energy spendings

1. **Prices vary with consumption →** estimate price of energy type n in country c in year t as a log linear function: $logP_{nct} = \alpha + \beta logQ_{nct} + \varepsilon_{nct}$
   – Raw data in quantity bands, e.g. in France 2019, price per TJ electricity was 51111 EUR if Q_{nct}<0.072TJ, but was 15916 EUR if Q_{nct}>540TJ



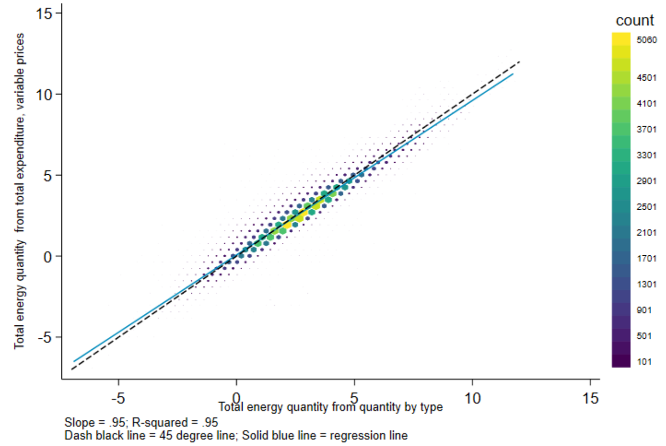
2. **Different energy types are used →** match firms to their industry k energy use shares: Q_{nkt}/Q_{kt}
   – E.g. French car manufacturing uses 57% electricity, 35% gas, etc
   – Some data at 2/3 digit level → use 4 digit French data to adjust proportionally
   – Large firms use lower share of electricity → use French data to estimate adjustment

3. **Varying total spending across firms →** estimate total quantity of energy Q used by each firm subject to their total spending, prices, and energy types

$$\max_q Q \ \ subject\ to$$

$$logP_n = \alpha_n + \beta_n logQ_n$$

$$S = Q \sum_n P_n \frac{Q_n}{Q}$$



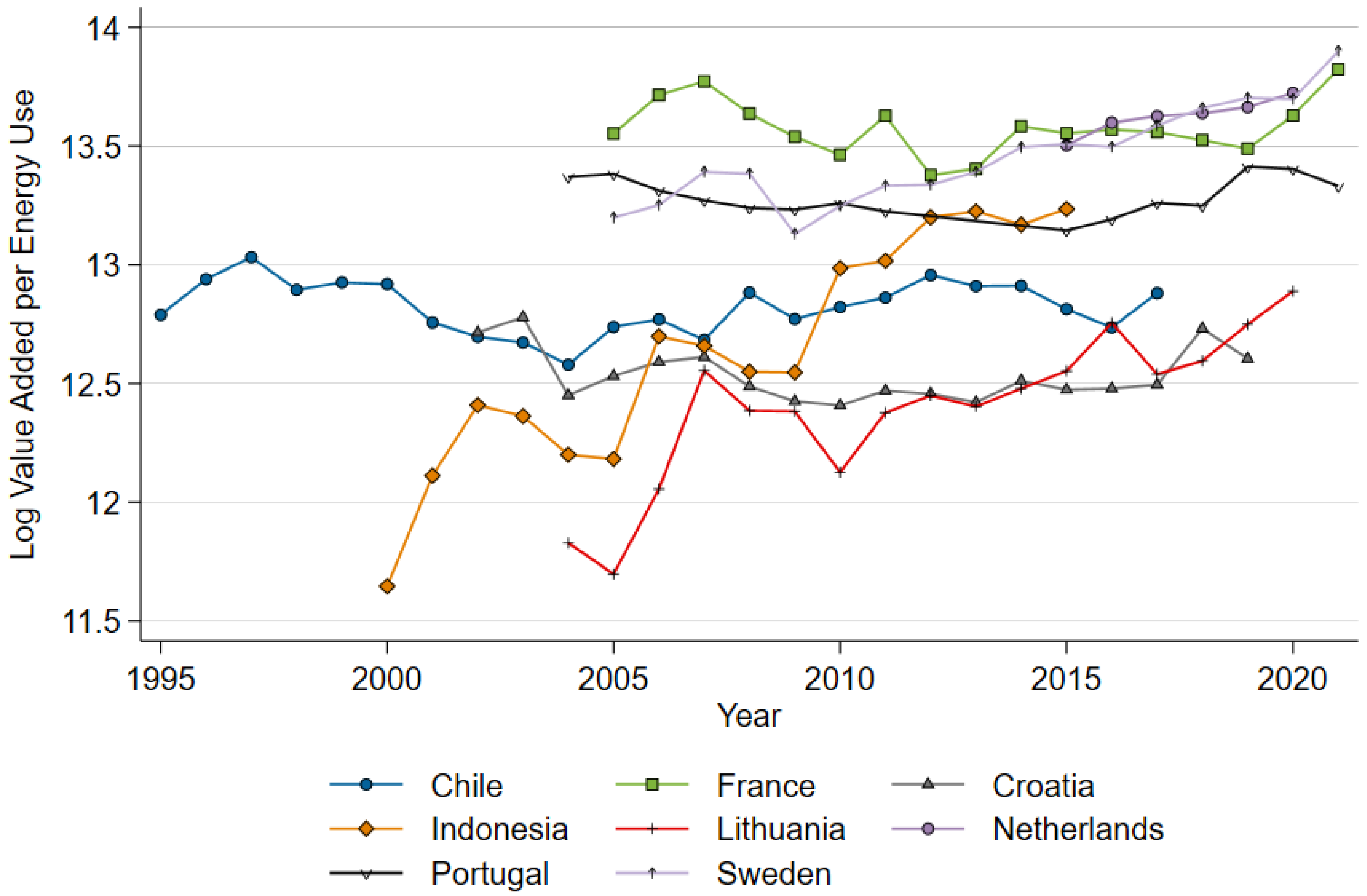
Firm-level summary stats for France

variable	mean	sd	p10	p50	p90	count
Employment	239.27	1989.92	26	79	432.02	105741
Value added (millions)	21.49	115	1.32	5.13	36.61	99049
CO2 Emissions (tonnes)	6455.63	103852.37	23.84	275.47	4756.76	105741
Log Value added / energy quantity	13.3	1.36	11.59	13.36	14.93	98130
Log Value added / CO2	9.89	1.52	7.98	9.95	11.72	98125
Log Value added / labour	11.1	0.59	10.47	11.09	11.77	98151
Firm age	30.01	21.11	8	26	54	99042
Energy spending / intermediate inputs	0.04	0.06	0	0.02	0.08	98652
Energy spending / total costs	0.02	0.04	0	0.01	0.05	98761
Intermediate input costs / total costs	0.7	0.14	0.52	0.72	0.87	98855
Labour costs / total costs	0.3	0.14	0.13	0.28	0.48	98855
Age of Capital	5.2	4.05	1	4	11	105741
Electricity quantity share (share of joules)	0.63	0.29	0.23	0.63	1	100986

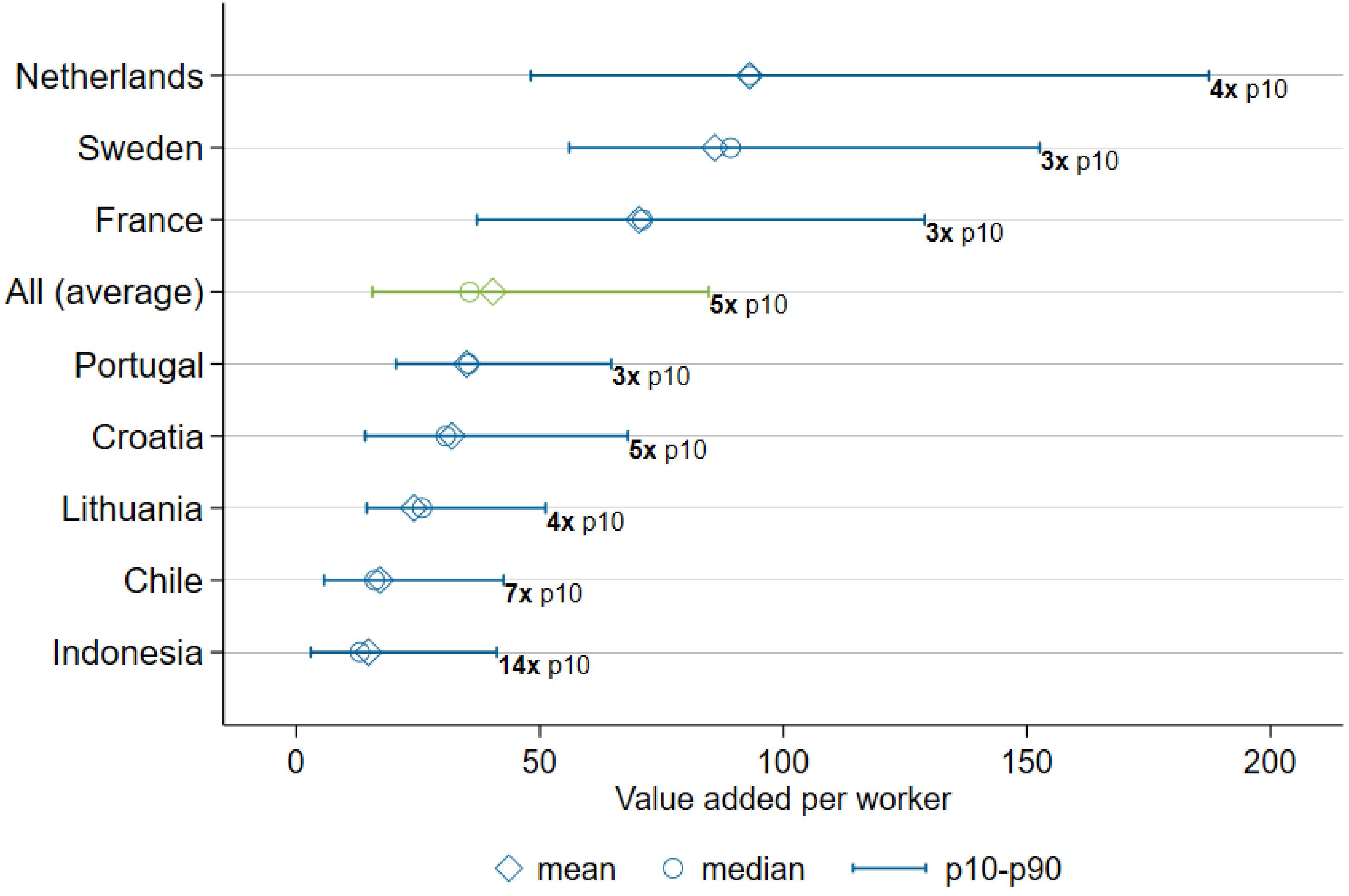
Country-industry-year summary stats

	mean	sd	min	p10	p25	p50	p75	p90	max	count
Aggregate statistics computed for all cells										
Firm Count	109	254	1	2	6	27	96	267	3278	8106
[sum] Employment	16184	37902	20	237	1027	4250	13610	39174	514142	7593
[sum] Value Added (millions)	733	1500	-735	4	32	171	673	2085	18003	7571
Firm Count (10 plus firms)	158	295	10	14	25	63	149	393	3278	5547
Dispersion statistics computed on cells with at least 10 firms										
[p10] Log Value added / energy	11.61	1.13	6.80	10.16	10.77	11.60	12.41	13.08	15.47	5166
[p50] Log Value added / energy	13.07	0.98	10.24	11.80	12.36	13.05	13.70	14.33	16.38	5166
[p90] Log Value added / energy	14.49	1.00	11.24	13.24	13.87	14.47	15.08	15.73	18.04	5166
[p10] Log Value added / CO_2	7.49	1.54	2.67	5.56	6.35	7.36	8.55	9.52	12.87	5165
[p50] Log Value added / CO_2	9.04	1.46	5.72	7.32	7.92	8.84	9.97	11.05	14.84	5165
[p90] Log Value added / CO_2	10.58	1.55	6.51	8.85	9.46	10.26	11.51	12.78	16.31	5165
[p10] Log Value added / labour	9.37	1.33	4.85	7.41	8.48	9.47	10.55	10.91	12.61	5307
[p50] Log Value added / labour	10.23	1.02	6.22	8.90	9.55	10.26	11.10	11.40	13.59	5307
[p90] Log Value added / labour	11.08	0.88	8.18	9.96	10.49	11.13	11.69	12.13	14.98	5307
Log(VA/energy) p90-p10	2.88	0.88	0.28	1.83	2.25	2.78	3.44	4.03	7.17	5166
log(VA/CO_2) p90-p10	3.09	0.91	0.27	1.96	2.43	3.05	3.66	4.30	7.13	5165
log(VA/L) p90-p10	1.71	0.88	0.29	0.84	1.06	1.43	2.18	3.07	5.21	5307

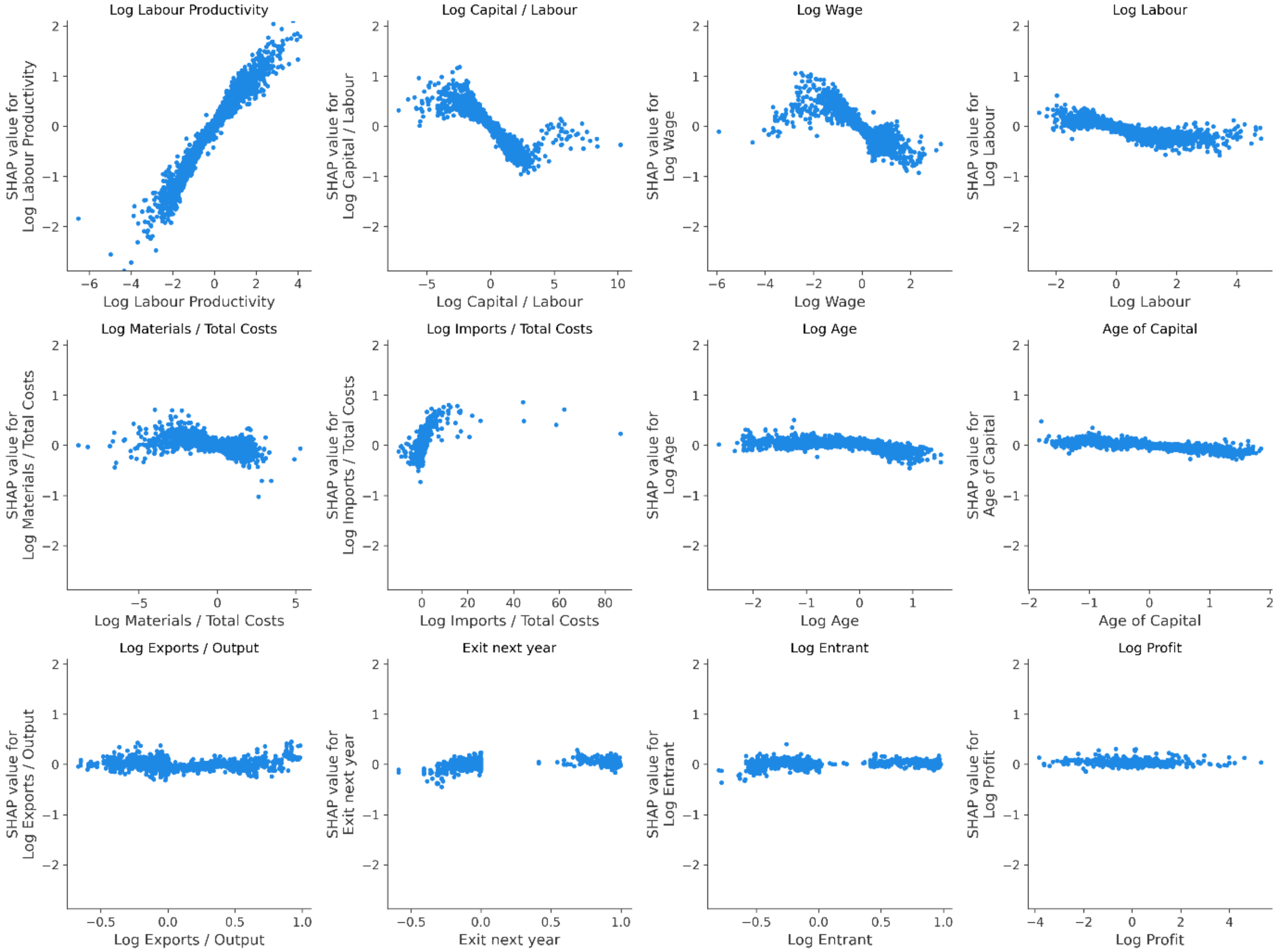
Average energy productivity has improved in most countries



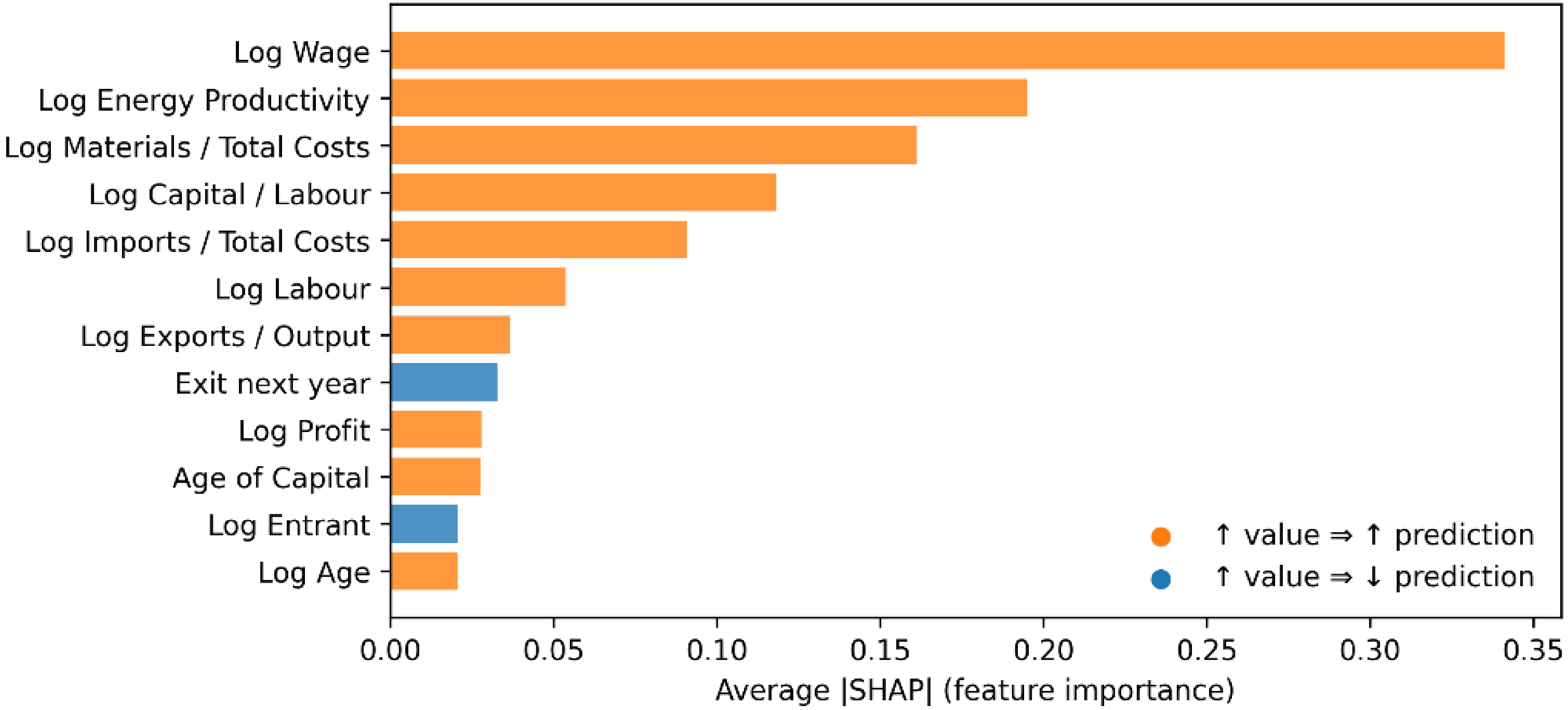
Percentiles and dispersion in labour productivity



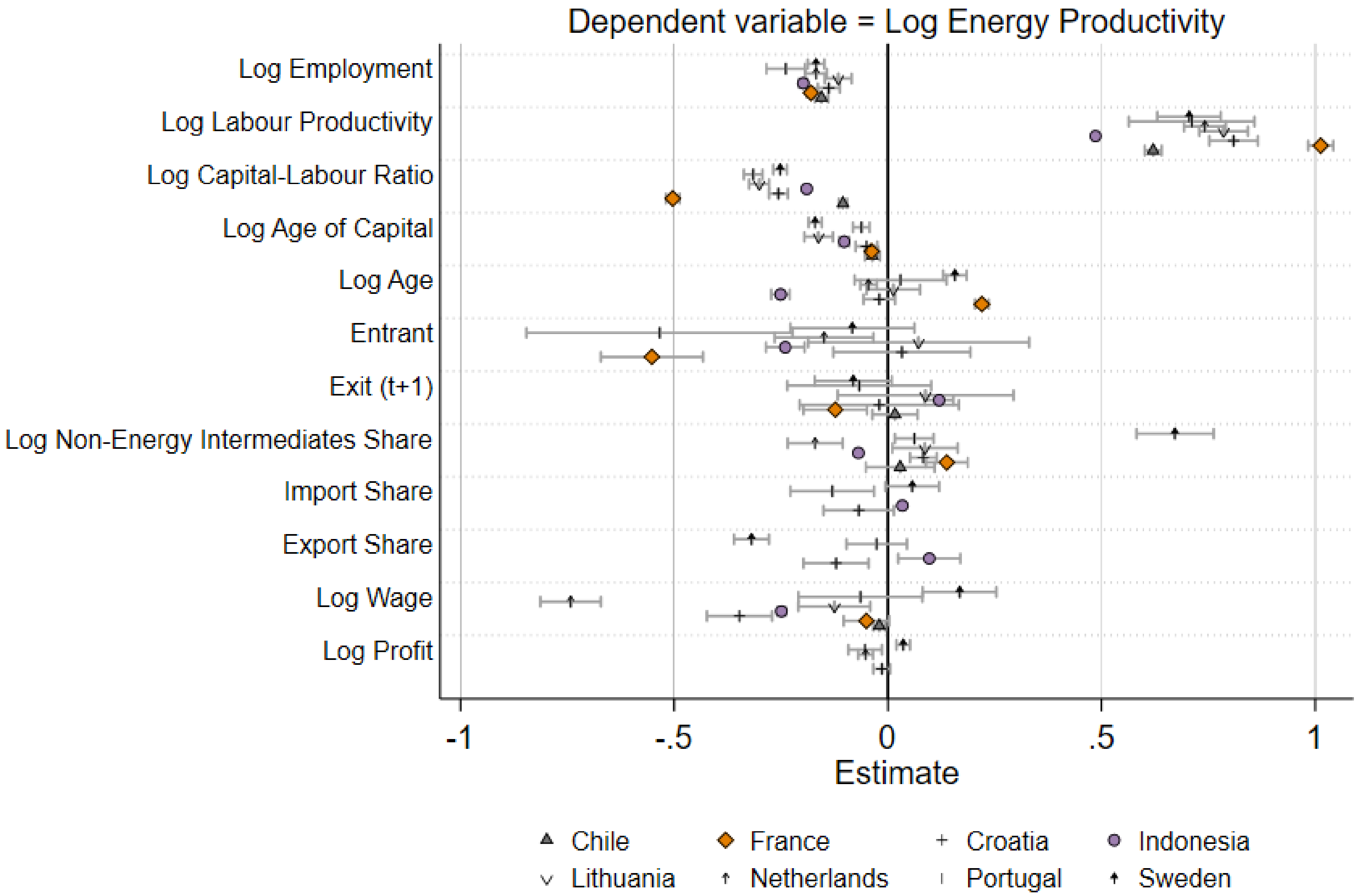
Relationship between SHAP and predictor values for firms' energy productivity



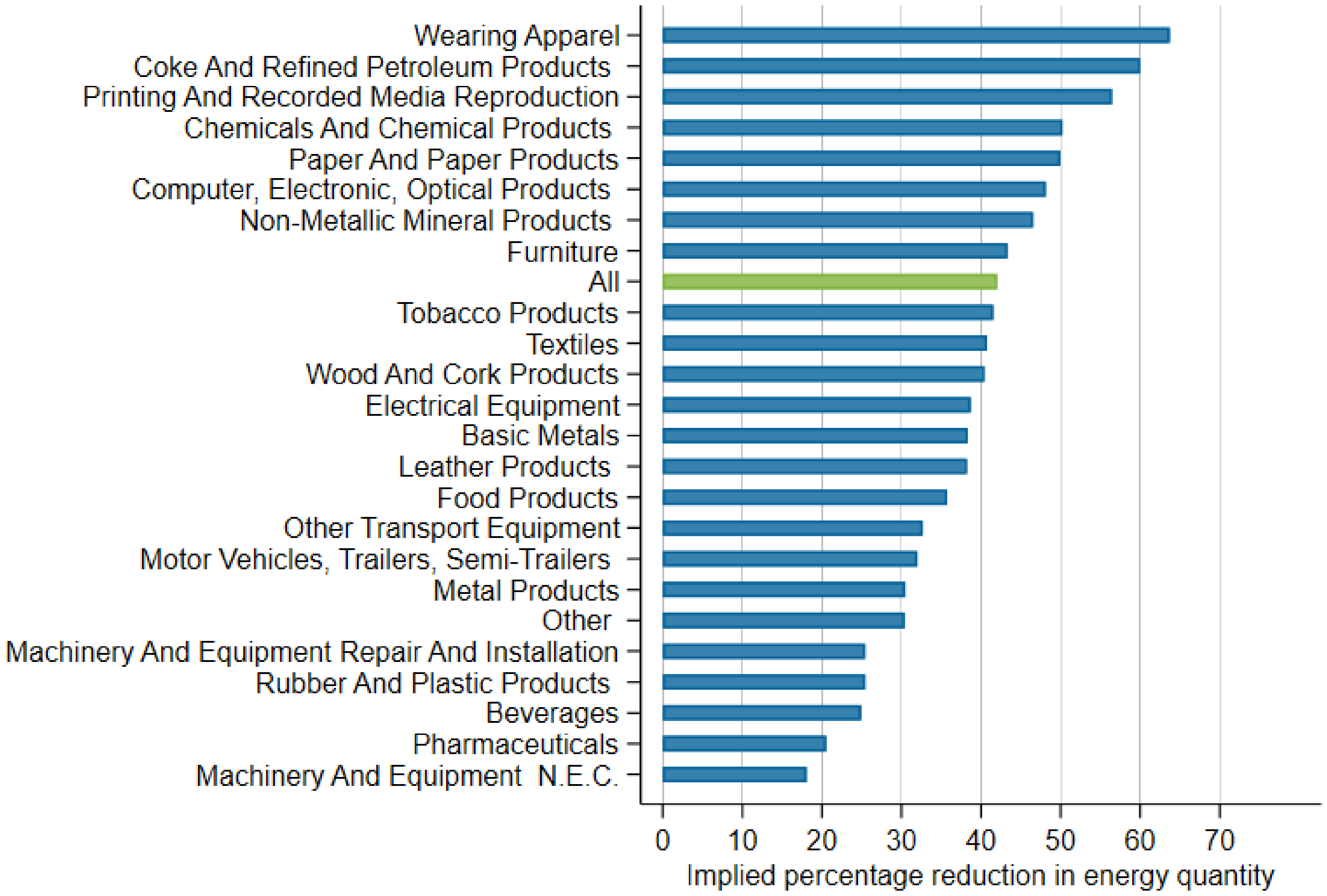
Labour productivity predictors



Multivariate regressions for firm energy productivity



Potential reductions in energy productivity by industry

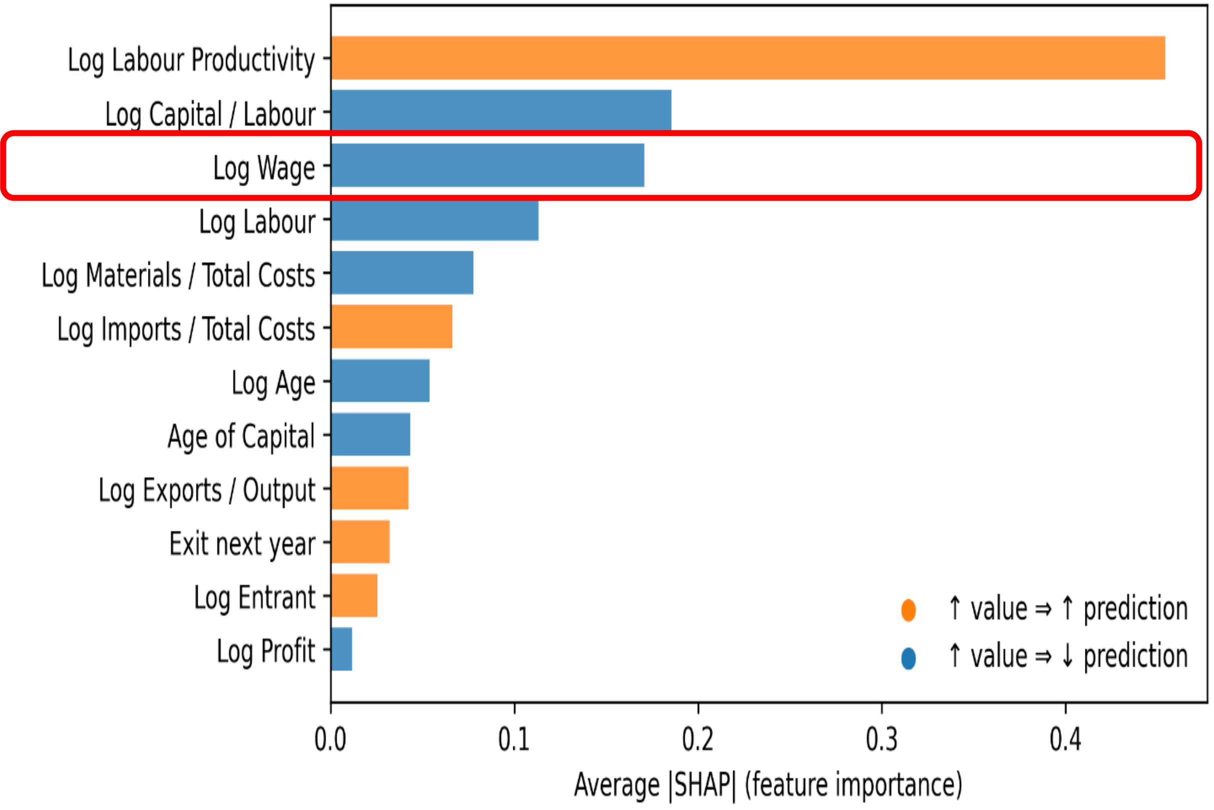


XGBoost prediction of energy productivity dispersion: model performance

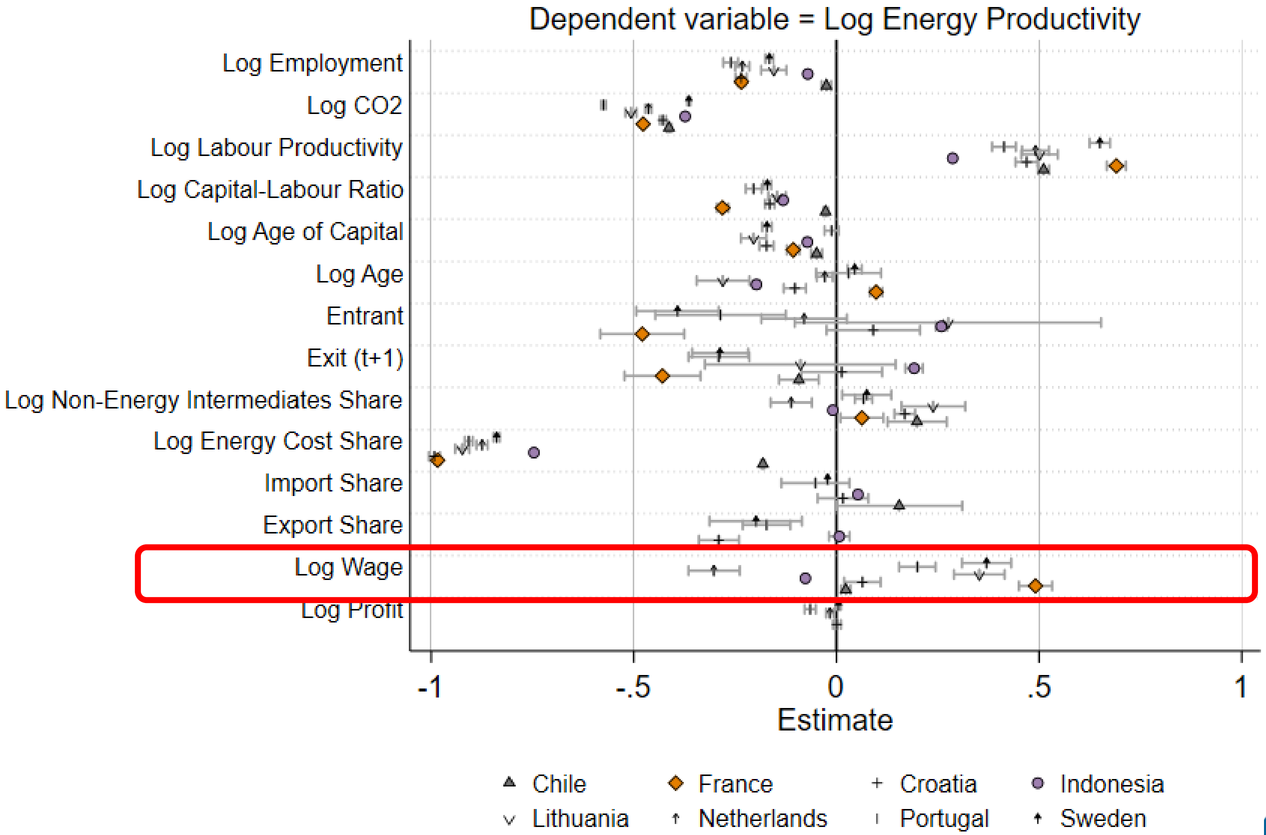
	XGBoost (test sample)	OLS (test sample)	OLS (full sample)
Observations	1 031	135	680
Root mean square error	0.53	0.61	0.54
R-squared	0.46	0.18	0.34

Controlling for other factors, firms paying higher wages have lower energy productivity

ML Prediction: Feature importance

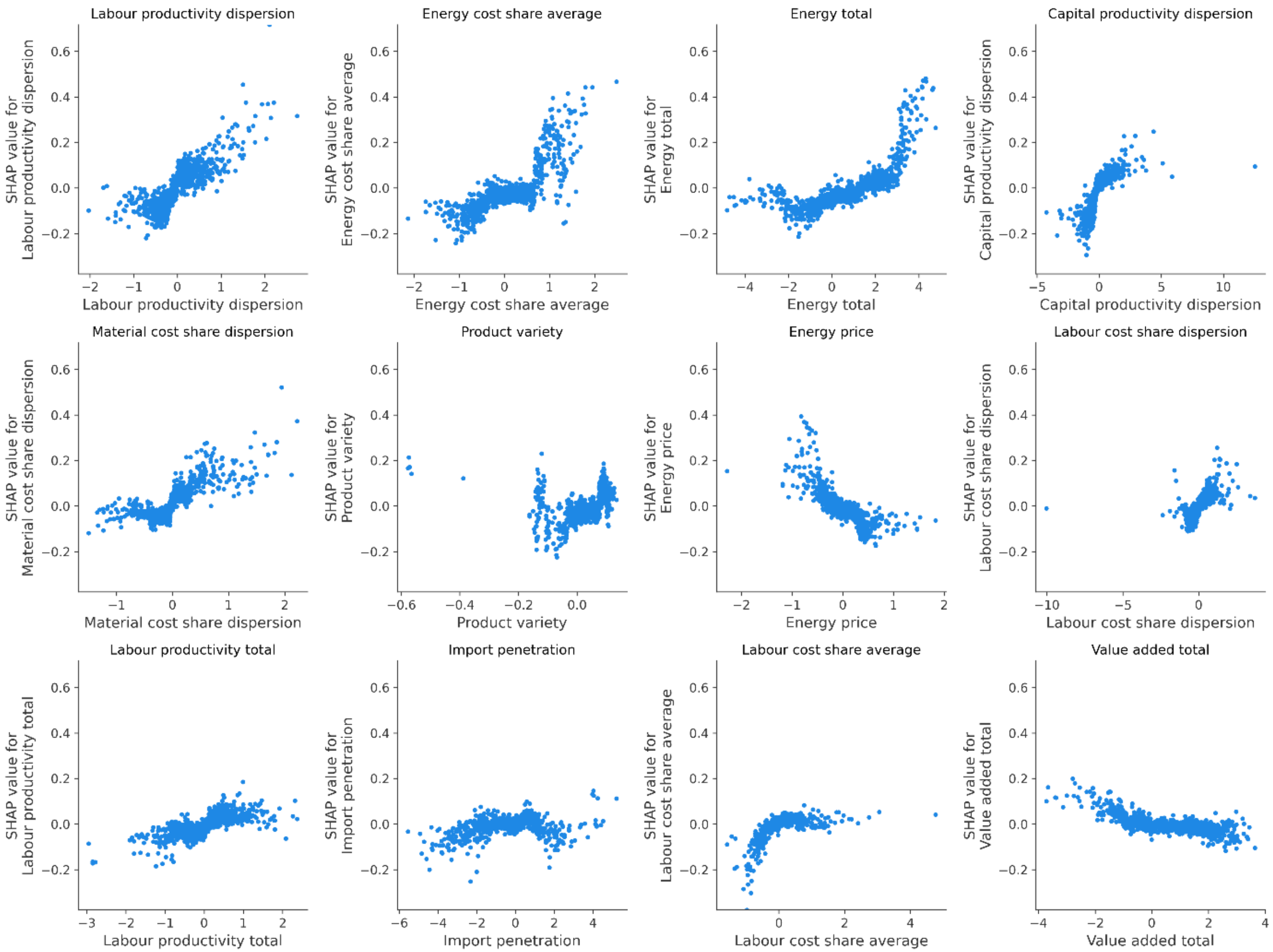


Bivariate OLS: estimated coefficients

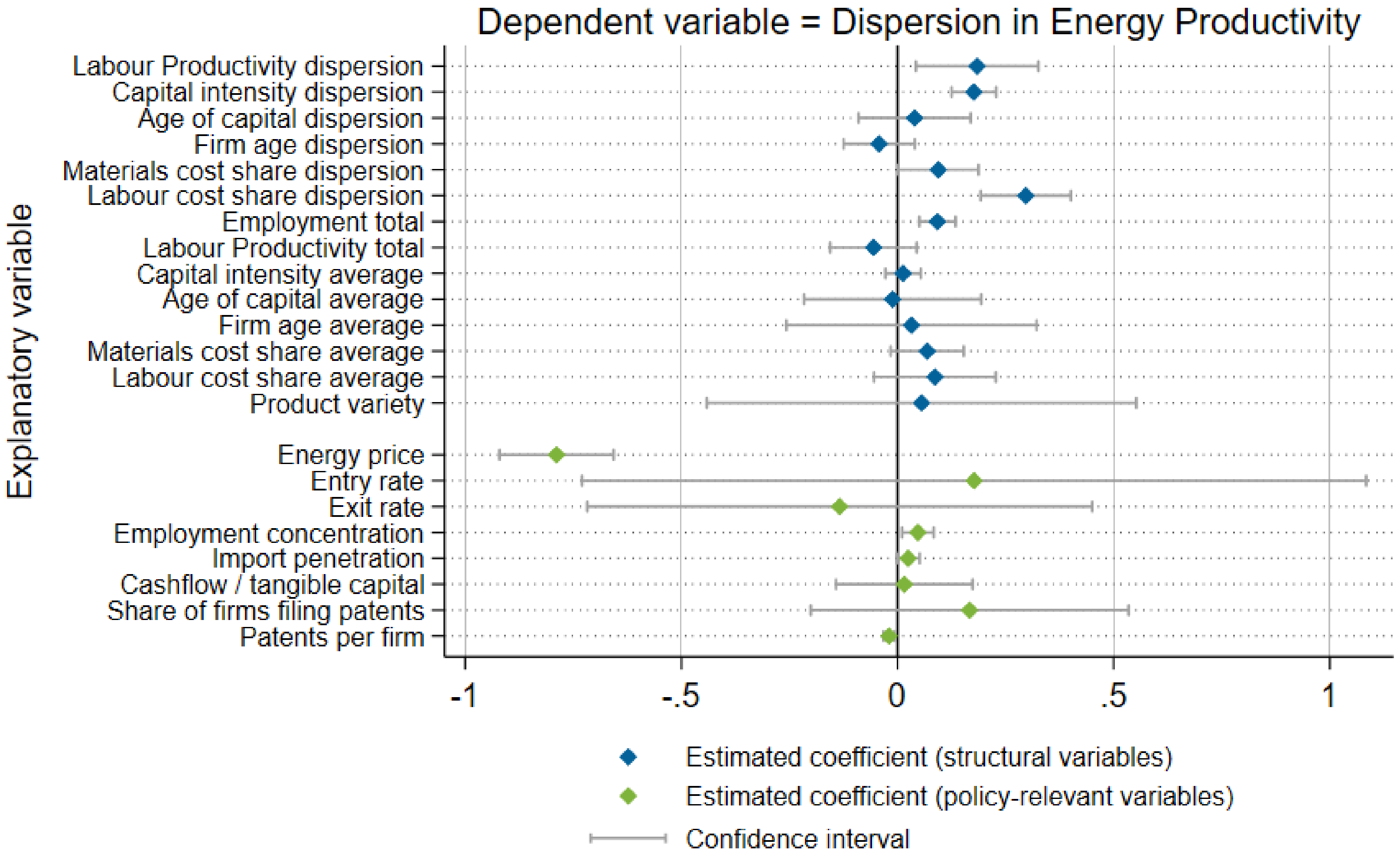


- With controls higher wages are associated with lower energy productivity
- Possible explanations? **Management practices?**

Relationship between SHAP and predictor values for industry dispersion in energy productivity



Multivariate OLS regressions for energy productivity dispersion



Observations: 1931
Croatia=341, France=573, Indonesia=457, Lithuania=69, Portugal=211, Sweden=280

5-year changes OLS regressions for energy productivity dispersion