Energy Security and Firm Adjustment: Employment and Energy Management after Europe's Gas Shock

Hazwan Haini[1]

Abstract

Does exposure to an energy-security shock coincide with organizational adaptation, reductions in employment, or both? This paper links World Bank Enterprise Survey establishments in a 2019 baseline cohort to post-crisis interviews and constructs a predetermined exposure index from countries' recorded Russian natural-gas import shares in 2020 and a gas-weighted U.S. manufacturing energy benchmark for 2018. First-difference regressions include country and baseline-industry fixed effects and baseline firm characteristics. Among 810 manufacturing establishments in nine European Union economies, greater exposure is not associated with a statistically distinguishable increase in reported energy-management measures. The coefficient is -0.097 (country-clustered standard error 0.071), corresponding to -3.95 percentage points for a one-standard-deviation increase in exposure. In a larger employment sample of 849 establishments, exposure is negatively associated with changes in log permanent full-time employment: -0.095 (0.022). Supplementary restricted wild-cluster inference, enumerating 512 country-level sign patterns, yields p-values of 0.234 and 0.0039, respectively. The employment finding remains negative in common-sample and leave-one-country-out checks. These patterns distinguish an observed employment-adjustment margin from evidence of accelerated energy management; they do not identify emissions reductions, permanent technology adoption, or a causal response to carbon pricing. The interpretation is limited by nine country clusters, selective reobservation, imperfect exposure measurement, overlapping recall periods, questionnaire changes, and the absence of a comparable pre-crisis energy-management trend.

Keywords: energy security; firm adjustment; energy management; natural gas; manufacturing; enterprise surveys.

JEL classification: D22; L60; Q41; Q55.

---

[1] Centre for Strategic and Policy Studies, Brunei Darussalam

Corresponding Author Email: hazwan.haini@csps.org.bn

1. Introduction

A large increase in the cost of fossil energy can change both how firms produce and how much they produce. The first response is potentially adaptive: firms may reorganize energy use, modify equipment, or improve monitoring. The second is a scale response: firms may reduce operations, employment, or investment. These margins have different implications for economic activity and environmental performance. A decline in energy demand does not, by itself, establish that production has become more efficient. Conversely, an efficiency improvement need not reduce total energy use when it also makes production more profitable. Distinguishing the margins is therefore central to interpreting firm responses to energy disruption.

Europe's gas crisis provides a setting in which this distinction matters. Russian pipeline supply reductions during 2022 interacted with an already tight energy market, affecting prices and prompting substantial adjustment in European gas demand (International Energy Agency, 2023a). The International Energy Agency estimates that EU gas consumption fell by 13% in 2022 and industrial consumption by approximately 25%. Its accounting attributes industrial reductions to several channels, including production curtailment, fuel switching, and other changes, rather than to efficiency alone (International Energy Agency, 2023b). Aggregate evidence, however, does not reveal whether establishments with greater prior exposure were more likely to change their management practices or to reduce employment.

This paper examines that question using a linked establishment panel from the World Bank Enterprise Surveys (WBES). The baseline is the survey cohort labelled 2019, and follow-up observations are drawn from the latest available post-crisis waves. The primary sample contains 810 manufacturing establishments in Bulgaria, Cyprus, Hungary, Italy, Malta, Poland, Portugal, the Slovak Republic, and Slovenia. Predetermined exposure combines each country's recorded Russian share of natural-gas imports in 2020 with a U.S. industry energy benchmark measured in 2018. The latter is a gas-weighted fuel-energy-intensity proxy, not a direct measurement of European establishments' gas requirements. All exposure assignments are fixed using baseline country and industry.

The empirical specification relates changes in establishment outcomes to this exposure index, controlling for country effects, baseline-industry effects, and baseline size, age, exporting, foreign ownership, and reported financing obstacles. Country effects account for changes common to establishments in the same country. Industry effects account for changes

common to the same manufacturing division across countries. Consequently, the coefficient captures whether the relationship between industry energy requirements and subsequent adjustment differs across countries with different recorded Russian import shares. It is not a simple comparison between EU and non-EU establishments, nor does it instrument a measured establishment-level energy price.

The results do not support an affirmative claim that greater exposure accelerated reported energy-management adoption. The primary coefficient is -0.097, with a country-clustered standard error of 0.071 and a conventional p-value of 0.210. The exposure index has a standard deviation of 0.407 in the primary sample. At that scale, the point estimate is -3.95 percentage points, with a conventional 95% interval extending from -10.63 to 2.73 percentage points. Controlling for the baseline outcome produces a coefficient of -0.065; applying baseline sampling weights produces an estimate close to zero. The absence of statistical significance is not evidence that all forms of adaptation were absent. It concerns one retrospective, self-reported organizational action and one exposure gradient.

A different pattern appears for permanent full-time employment. Among 849 establishments with the requisite employment observations and controls, the coefficient on exposure in a log-employment-change regression is -0.095, with a standard error of 0.022. This association remains negative when employment is estimated on the 810-establishment energy-management sample, when baseline log employment is omitted, and when individual countries are removed. Supplementary restricted wild-cluster tests yield p-values of 0.234 for energy management and 0.0039 for employment. The latter becomes 0.027 after adjusting the seven primary and secondary outcome tests using Holm's procedure. Employment is nevertheless a secondary outcome, and neither inferential adjustment establishes that the gas shock caused the observed changes.

The paper contributes by placing organizational response and employment adjustment within the same exposure design, rather than inferring adaptation from aggregate energy savings. It also makes the measurement content of the design explicit. The survey outcome records whether an establishment reports measures introduced during the preceding three years, not the stock of energy-management capability or its physical effectiveness. The sector benchmark mixes fuel-intensity and first-use energy statistics and is therefore interpreted as a proxy. The cross-country panel adds evidence from establishments and economies that are distinct from the German and U.S. firm evidence studied by Amaglobeli et al. (2024), while retaining the limitations of a relatively small and selectively observed survey panel.

The analysis is most appropriately read as a conditional exposure-gradient study. A causal interpretation requires the absence of unobserved country-industry developments correlated with exposure, stable outcome measurement, and sufficiently benign sample selection. The data cannot verify those conditions. In particular, the change interval spans pandemic-related disruption and recovery, the primary outcome has no comparable earlier pre-period, and actual post-crisis interviews occur at different dates. These limitations are substantive features of the research design, not matters that additional fixed effects or a larger number of firm records automatically resolve.

2. Economic context and related research

2.1. Energy security, prices, and the margins of adjustment

The 2022 crisis involved a physical supply disruption, changes in expectations, and responses in interconnected energy markets. Such an episode is not equivalent to a uniform, permanent increase in an administratively set carbon price. Firms can differ in contracting arrangements, access to substitutes, indirect electricity exposure, and the timing with which input costs reach them. The IEA's contemporaneous assessment emphasizes both lower Russian pipeline deliveries and adjustment through liquefied natural gas imports, storage, and reduced consumption (International Energy Agency, 2023a). These market-wide responses imply that a low recorded Russian import share does not identify a country insulated from the crisis.

The macroeconomic consequences also extend beyond a firm's direct gas bill. An adverse energy terms-of-trade shock can change demand for domestic products and the allocation of resources between households and firms. Auclert et al. (2023) analyze such interactions in a model of fiscal and monetary responses to an energy shock. The present study does not estimate those policy mechanisms. The relevance of that framework is narrower: an establishment may face both higher input costs and changing demand, making a pure technology-response interpretation of employment or investment particularly demanding.

Evidence on firms during the crisis similarly points to multiple adjustment margins. Amaglobeli et al. (2024) investigate firm resilience and responses to fiscal incentives using evidence that includes Germany and the United States. Their analysis discusses investment and operational responses to energy pressures. The panel examined here differs in geography, sample composition, timing, and measurement. A weak energy-management exposure gradient in these nine economies would therefore not contradict evidence of efficiency

investment among German establishments. It addresses a different outcome and population, over a different observation window.

2.2. Induced innovation and organizational capability

The induced-innovation literature explains why relative energy prices can redirect technological effort. Popp (2002) links historical U.S. energy-related innovation to energy prices and the existing knowledge stock. Aghion et al. (2016) study directed technical change in the automobile industry and emphasize both fuel-price incentives and technological path dependence. These mechanisms motivate a possible adaptive response, but they do not imply that every short-run supply disruption generates observable organizational investment. The expected persistence of prices, available technologies, and the cost of adjustment all matter.

A separate literature concerns the organization of energy use inside firms. Bloom et al. (2010) associate better management with lower energy intensity, while Martin et al. (2012) examine relationships among organizational structure, management practices, and energy efficiency. These studies motivate energy management as an economically meaningful organizational margin. They do not make a binary retrospective survey response equivalent to a measured efficiency improvement. In particular, general managerial capability and a report of a recent energy-management measure are related concepts but different empirical objects.

The energy-efficiency-gap debate further cautions against reading every unadopted measure as a missed profitable opportunity. Allcott and Greenstone (2012) discuss the empirical challenges in identifying privately worthwhile efficiency investments and the possible roles of market and behavioral failures. For this paper, the relevant distinction is between the existence of a potentially profitable technology and its adoption by an establishment with particular scale, financing, and information constraints. The observed data do not identify the net present value of an unadopted measure, so non-adoption cannot establish inefficiency.

Finally, adaptation can alter scale rather than simply conserve energy. In an experiment offering energy consulting to Indian manufacturing plants, Ryan (2018) finds changes in operating hours and skilled-labor demand, illustrating complementarity between energy productivity and other inputs. The implication is not that the European establishments should behave similarly. It is that energy use, employment, and efficiency need not move in the same direction. A research design that measures only one of these outcomes cannot recover the others without additional assumptions.

3. A framework for adaptation and employment adjustment

Consider an establishment that chooses output, labor, and an energy-management action. Let q_i denote output, L_i labor, p_i the effective price of energy, and a_i energy required per unit of output before the action. A management measure lowers this requirement by a proportion θ_i, with $0 < \theta_i < 1$, and entails a fixed implementation cost K_i. It may also require scarce managerial time. The establishment evaluates the investment using expected energy prices, output, and a discount rate that can reflect financing conditions.

For a fixed expected output path, a stylized present value of the action is

$$V_i = \sum_{s=0}^{T} \frac{\mathbb{E}_i[p_{is}\theta_i a_i q_{is}]}{(1+r_i)^s} - K_i. \tag{1}$$

An establishment adopts when this value exceeds the relevant implementation and financing threshold. Equation (1) is an organizing device, not an estimated structural model. It abstracts from equipment replacement, learning, and interactions with other inputs. Its usefulness is that the economic value of conserving energy depends jointly on the price of energy and the scale on which the establishment expects to operate.

If output were fixed, a higher expected energy price would raise the gross savings from a given measure. Once output responds, that conclusion need not follow. In a one-period version with $q_i = q_i(p_i)$, the derivative of gross savings is

$$\frac{d(p_i\theta_i a_i q_i(p_i))}{dp_i} = \theta_i a_i\{q_i(p_i) + p_i q_i'(p_i)\}. \tag{2}$$

The expression can be small or negative when output falls sufficiently strongly. Thus, even before introducing financing constraints or uncertainty, a cost shock can reduce the scale over which an efficiency measure would pay for itself. Higher financing costs or a shorter perceived duration of the price increase can further weaken investment. None of these channels is separately identified by the survey regressions; they explain why a positive organizational response is a hypothesis rather than an accounting necessity.

Labor adjustment is equally conditional. With labor and energy complementary in the relevant production range, lower desired output can reduce employment. Yet a successful organizational or technological change may increase productivity and output, potentially supporting employment. The combination of a weak management-action gradient and a negative employment gradient is therefore consistent with adjustment concentrated on operating scale, but it does not demonstrate that management investment and labor were

substitutes. Establishments may use unmeasured efficiency practices, and both measured outcomes may respond to a third shock.

The empirical outcome adds another distinction. Let A_{it} indicate that an establishment reports introducing a measure during the three years preceding interview t. The observed change is $A_{i1} - A_{i0}$, not the change in an underlying stock of capability. A firm reporting an action at baseline but none at follow-up may still retain and benefit from the earlier measure. Likewise, reporting an action in both periods may reflect repeated improvements rather than a persistent adoption state. The analysis consequently tests whether exposure is associated with changes in the incidence of recent action, without assigning a technological depreciation interpretation to transitions between survey responses.

4. Data, measurement, and sample construction

4.1. Establishment panel and sample restrictions

The WBES collect information on private establishments, including performance, employment, ownership, trade, business constraints, and selected environmental practices. Standard surveys cover formal establishments with at least five workers in eligible industries and employ stratified sampling. The methodology also provides for reinterviewing establishments from earlier rounds (World Bank, n.d.). An establishment, rather than a consolidated multinational enterprise, is the statistical unit. The term firm is used below as shorthand for that unit. The primary sample includes establishments reporting as few as three permanent full-time workers; this employment measure and the reinterviewed sample should not be conflated with the standard sampling-frame threshold.

The consolidated data contain 22,099 firm-wave observations across the retained historical waves of 14 economies. There are 7,609 establishments in the 2019-labelled baseline cohort, of which 2,179 are observed in the selected later wave. Valid energy-management answers at both dates are available for 2,053 establishments. The panel identifiers are formed within country before linking waves; country-specific identifiers are not assumed to be globally unique.

The exposure analysis is more restricted than this broad descriptive sample. Consistent country-level Russian import shares are included for nine EU economies, and the sector benchmark covers manufacturing divisions that can be mapped to the U.S. energy data. Among establishments with energy-management pairs, 834 have both exposure components.

Requiring all baseline controls leaves 810 establishments, occupying 129 country-industry cells across 20 manufacturing divisions. Twenty-four exposure-ready pairs are lost at this final complete-case step. The remaining five economies are retained for descriptive comparisons, not assigned artificial zero exposure.

Table 1. Country coverage and the primary manufacturing sample

Country	Baseline N	Reobserved N	EM pairs	Main N	Russian share (%)	Post wave
Bulgaria	772	218	198	103	75	2023
Cyprus	240	80	76	27	0	2024
Hungary	805	295	290	170	95	2023
Italy	760	56	56	34	43	2024
Malta	242	95	95	31	0	2024
Poland	1369	185	154	98	55	2024
Portugal	1062	308	302	215	10	2023
Slovak Republic	429	125	111	60	85	2024
Slovenia	409	196	184	72	9	2024
Total	6088	1558	1466	810		

Notes: Baseline, reobservation and energy-management (EM) pair counts include all retained sectors in each country; Main N applies manufacturing exposure and complete-control restrictions. Russian share is the recorded 2020 share of Russian gas in total natural-gas imports. Zero shares do not imply no energy-crisis exposure. Source: WBES panel and European Commission (2022).

The primary sample is uneven across countries. Portugal contributes 215 establishments and Hungary 170, while Cyprus contributes 27 and Malta 31. This imbalance matters for both the pooled estimand and inference. The nine countries are neither a census of EU manufacturing nor a random sample of European energy systems. Sample selection is determined by the available survey waves, successful matching, item response, and exposure coverage, rather than by a claim to represent all establishments affected by the crisis.

4.2. Energy management and the timing of measurement

The baseline outcome uses the energy-management item in the 2019 BEEPS environmental module, recorded as BMGc23d. The corresponding later variable is ge8d. Both concern actions introduced during the preceding three years, but the question formats are not identical. The baseline asks about energy management as one item in a list of measures; the 2023 standard instrument asks a standalone energy-management question with an explicit emissions, waste, or pollution purpose (World Bank, 2019, 2023). Positive and negative responses are coded one and zero; nonresponse and inapplicable codes are not recoded as zero.

This harmonization establishes conceptual overlap, not measurement invariance. Differences in framing, questionnaire context, or respondent interpretation can affect levels and possibly changes. Country and industry effects remove only suitably additive differences, not exposure-related changes in reporting. The analysis therefore uses the label reported energy-management measures throughout. It does not equate the outcome with audited energy savings, ISO certification, a carbon-emissions reduction, or a permanent change in organizational capability.

Survey labels must also be distinguished from actual fieldwork. Primary-sample baseline interviews occurred during 2018-2020, despite the common 2019 cohort label. Post interviews occurred during 2023-2025. The Slovak Republic's later wave is labelled 2024 in the panel but its primary-sample interviews took place in 2023. The average interview interval is 4.50 years. Follow-up three-year recall windows can include periods before the 2022 disruption, and the precise adoption dates are unknown. Only 57 primary-sample observations, from Malta and Poland, pass the available year-based flag for a recall period beginning no earlier than 2022. Even that flag does not isolate the period after the February invasion. Appendix B reports timing explicitly.

Employment and financial variables have a different time reference: the relevant completed fiscal year. Consequently, log-employment changes span the establishment's observed fiscal-year endpoints rather than a common calendar-year treatment window. They should not be read as employment losses occurring exclusively in 2022 or as annual growth rates. The longer interval includes both the pandemic and its aftermath. This is an important source of potentially confounding country-industry variation.

4.3. Predetermined country and sector exposure

The country component, R_c, is the recorded share of Russian imports in total natural-gas imports in 2020, drawn from the European Commission's country Energy Snapshots (European Commission, 2022). The supplied shares range from zero in Cyprus and Malta to 0.95 in Hungary. They describe gross recorded import sourcing, not the Russian share of domestic gas consumption. Domestic production, reexports, indirect Russian-origin gas traded through other countries, and market integration complicate the link between this measure and the prices paid by establishments. Cyprus is explicitly recorded as having no gas use in the source, making its zero substantively different from diversified gas importing.

The sector component is constructed outside the European sample using the 2018 U.S. Manufacturing Energy Consumption Survey. Denote fuel consumption per dollar of shipments for NAICS industry k by f_k, total first-use energy for fuel and nonfuel purposes by E_k^{all}, and natural gas on that same first-use basis by G_k^{all}. The recorded benchmark is

$$J_k = f_k \frac{G_k^{all}}{E_k^{all}}. \tag{3}$$

Fuel intensity comes from MECS Table 6.1, while the energy mix comes from Table 1.2 (U.S. Energy Information Administration, 2021b, 2021a). The distinction between the tables is material. First-use energy includes nonfuel uses, whereas the intensity statistic measures fuel consumption. Their product is therefore a gas-weighted fuel-energy-intensity proxy. It is not an identity for gas consumption per dollar of shipments. The manuscript preserves the benchmark values used in estimation but does not retain the stronger interpretation that the formula cannot support.

For baseline ISIC divisions mapped to multiple NAICS categories, the existing aggregation rule is

$$J_j = \frac{\sum_{k \in \mathcal{K}(j)} G_k^{all}}{\sum_{k \in \mathcal{K}(j)} E_k^{all} / f_k}. \tag{4}$$

The denominator in (4) is a constructed aggregation quantity, not observed shipments, because the energy accounting boundaries differ. Appendix A lists all mappings and resulting values. Several mappings combine broad industries; others assign the same U.S. category to more than one ISIC division. These choices preserve feasibility but introduce uncertainty about the correspondence between U.S. benchmark technology and European establishment activities.

Exposure is the product

$$Z_{cj} = R_c J_j. \tag{5}$$

The two components precede the crisis and are fixed at baseline country and industry. Predetermination avoids using a sector energy mix already changed by the shock, but it does not make the product exogenous. Country-industry specialization, energy substitution possibilities, and supply-chain exposure may still correlate with subsequent outcomes. Measurement error need not be classical or merely attenuating. The coefficient is reported per index unit, with a common standard-deviation translation for interpretation, rather than as a response to a known percentage increase in gas prices.

4.4. Outcomes, controls, and descriptive statistics

Permanent full-time employment is logged before differencing, with nonpositive or invalid observations excluded. Secondary outcomes are capacity utilization, electricity expenditure divided by sales, an outage indicator, generator ownership or sharing, and exporter status. Electricity expenditure relative to sales is a nominal spending share, not physical electricity intensity; ratios exceeding 100% are excluded by the original cleaning rule rather than winsorized. No deflator is supplied in the executed analysis, so real sales, real sales per worker, and total factor productivity are not estimated.

Controls are measured at baseline: log permanent full-time employment, age calculated from interview year and the reported year operations began, positive direct or indirect exports, foreign ownership of at least 10%, and a major or very severe financing obstacle. The size interaction uses fewer than 100 baseline employees, including smaller panel establishments; it is not the EU's legal SME definition. Holding classifications fixed prevents contemporaneous exporting, ownership, financing, or industry choices from being mechanically treated as predetermined controls.

Table 2. Primary-sample descriptive statistics

Variable	Mean	SD	Minimum	Maximum
Energy-management action, baseline	0.412	0.493	0.000	1.000
Energy-management action, follow-up	0.547	0.498	0.000	1.000
Change in energy-management action	0.135	0.652	-1.000	1.000
Permanent full-time employees, baseline	106.63	247.82	3.00	3500.00
Establishment age, baseline (years)	27.00	19.41	2.00	202.00
Exporter, baseline	0.575	0.495	0.000	1.000
Foreign ownership >=10%, baseline	0.109	0.311	0.000	1.000
Major/severe financing obstacle, baseline	0.083	0.276	0.000	1.000
Fewer than 100 employees, baseline	0.726	0.446	0.000	1.000
Predetermined exposure index	0.285	0.407	0.000	2.712
Interview interval (years)	4.50	0.73	3.08	6.08

Notes: N = 810 for every row. Unweighted sample statistics. Binary variables take values zero and one. Employment and age are measured at baseline. Exposure is the country import-share component multiplied by the sector proxy; its units are not a measured gas-price change.

In the primary sample, the fraction reporting energy-management action rises from 41.2% to 54.7%. There are 234 establishments moving from no reported action to reported action and 125 moving in the opposite direction; 209 report action in both periods and 242 in neither. The average establishment has approximately 107 permanent full-time employees,

although the median is 30.5. About 57.5% export and 10.9% have at least 10% foreign ownership. These are unweighted characteristics of the complete-case analytical sample, not population estimates.

4.5. Reobservation and external validity

Only 28.6% of the broad baseline cohort appears in the selected post wave. This is a reobservation rate, not a business-survival rate. It combines decisions about panel sampling, successful contact, eligibility, response, and any actual exit. Italy's broad reobservation rate is 7.4%, compared with 58.0% in Montenegro, illustrating why mechanically classifying absent establishments as closures would be inappropriate. The econometric sample conditions on successful reobservation and outcome availability.

Observable selection is present. In the broad baseline comparison, reobserved establishments more often report energy-management measures and exporting than establishments not reobserved. These differences do not identify the direction of bias in the exposure coefficient, but they rule out a casual assumption that retention is unrelated to business characteristics. Applying original survey weights changes the weighting of observed establishments; it does not recover the missing outcomes of establishments that were not reinterviewed. Any population or economy-wide interpretation therefore requires additional information about the sampling and response process.

5. Empirical strategy

5.1. First-difference specification and estimand

For establishment i, baseline country c, and baseline industry j, the primary specification is

$$\Delta Y_{icj} = \alpha_c + \gamma_j + \beta Z_{cj} + X_{i0}'\theta + \varepsilon_{icj}, \quad (6)$$

where $\Delta Y_{icj} = Y_{i1} - Y_{i0}$, X_{i0} contains the baseline controls, and α_c and γ_j are country and industry fixed effects in changes. Differencing removes time-invariant establishment-level outcome components. Country effects in (6) absorb country-specific changes common across included industries; industry effects absorb common changes for that baseline industry across countries. With only two analytical observations per establishment, the design cannot estimate an event-time trajectory.

For the management outcome, (6) is a linear model of a change taking values -1, 0, and 1. The linear specification provides an interpretable conditional gradient without assuming that

the change itself is a binary probability. The coefficient scaled by 100 describes percentage-point differences in the probability-change metric per exposure unit. A specification including Y_{i0} is reported separately. Algebraically, it is related to a post-outcome regression controlling for baseline outcome, and it can be informative about baseline differences; it is not an independent natural experiment.

The primary estimand weights observed complete-case establishments equally. A sensitivity specification uses baseline WBES median-assumption sampling weights. Those weights can change the implicit distribution of countries and establishment types in the pooled regression. Their use does not, without a complete response model, turn the analysis into a population-average causal effect. The distinction is consequential because the weighted and unweighted energy-management point estimates differ materially.

After removing country and industry fixed effects, the standard deviation of exposure is 0.235, compared with 0.407 before that projection. Thus, exposure is not mechanically collinear with the fixed effects. The presence of residual variation, however, establishes only that the coefficient is estimable. It does not establish that the variation is independent of the counterfactual evolution of outcomes.

5.2. Identification assumptions and competing explanations

A causal interpretation would require a conditional restriction of the form

$$\mathbb{E}\left[\Delta Y_{icj}(0) \mid Z_{cj}, c, j, X_{i0}\right] = \alpha_c + \gamma_j + X_{i0}'\theta, \tag{7}$$

where $\Delta Y(0)$ is the change that would have occurred without the crisis-related exposure channel of interest. Equation (7) excludes an unobserved country-industry trend correlated with $R_c J_j$, after conditioning on the specified effects and controls. In this setting, it is strong. Country-specific recovery in particular manufacturing industries, industry-targeted support, exposure to non-energy trade disruption, and differential product demand can all violate it.

The design also needs a sufficiently stable mapping from the observed index to the relevant economic exposure. Integrated gas and electricity markets create spillovers across countries, while contracts and substitution opportunities create differences within country-industry cells. A zero Russian import share is not a no-shock state. The estimated gradient consequently cannot be interpreted as the total effect of the European crisis, even under favorable assumptions about confounding. It is a comparison along one imperfectly observed dimension of predisposition to the shock.

There is no comparable pre-crisis series for the management-action outcome in the harmonized analysis. Earlier employment observations exist, but only five countries with the required exposure component contribute to the 2013-2019 diagnostic. The executed workflow does not estimate the exposure pretrend regression because this falls below its minimum country-cluster requirement. The available country-level changes do not substitute for a test of parallel counterfactual trends in the primary outcome. The paper therefore does not describe pretrends as validated.

Exposure shares common country and industry components. The broader literature shows why shared exposure structures can generate dependence across observations and invalidate naive firm-level inference (Adão et al., 2019). The present product is not a design with a large collection of independent shocks; results from many-shock asymptotics cannot simply be imported. Country clustering addresses within-country dependence but not all cross-country correlation within industries or through common energy markets. These considerations remain relevant even when the number of establishment observations exceeds 800.

5.3. Inference and supplementary calculations

The original estimates use country-clustered standard errors with finite-sample correction and conventional reference distributions based on eight degrees of freedom. With nine uneven clusters, the reliability of such inference depends on more than the total number of observations (MacKinnon et al., 2023). Confidence intervals and exact reported p-values are therefore shown rather than relying on significance stars alone.

Supplementary inference imposes the null value of the exposure coefficient and applies restricted wild-cluster bootstrap-t tests using Rademacher weights at country level. All $2^9 = 512$ sign patterns are enumerated. This removes simulation error for the chosen bootstrap distribution; it does not make the test an exact finite-sample randomization test or eliminate the assumptions about clustering. The procedure follows the restricted wild-bootstrap logic discussed by Roodman et al. (2019). The bootstrap calculations are additional to the executed Stata output, in which the optional bootstrap was switched off. Appendix C describes the independent coefficient validation and algorithm.

The seven primary and secondary outcome tests are also treated as one exploratory family for a supplementary Holm adjustment (Holm, 1979). This is a transparency check on multiple outcome inspection, not a claim that the study was preregistered. Heterogeneity models and

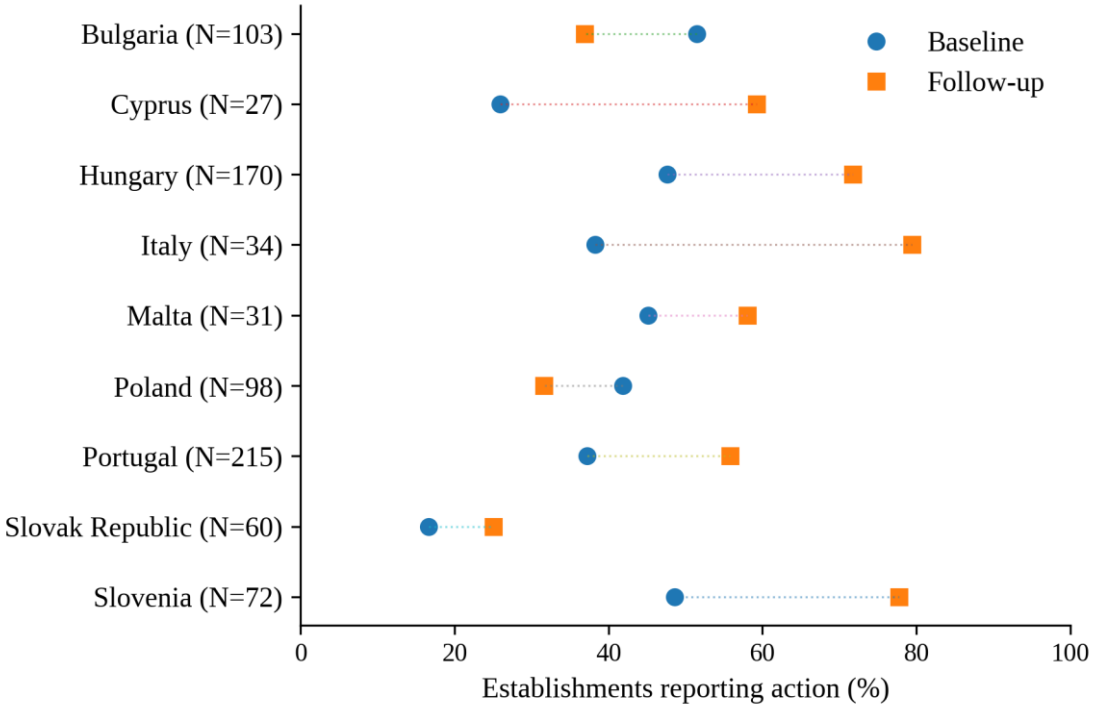
country exclusions remain exploratory, are reported as such, and are not used to redefine the primary hypothesis after viewing the results.

6. Results

6.1. Average changes do not identify an exposure response

Energy-management action becomes more common in the primary sample overall, but this aggregate increase is not the estimated exposure effect. Figure 1 displays substantial cross-country variation in the baseline and follow-up shares among the 810 complete cases. These means combine the effects of country conditions, industry composition, reporting, and selection. They are useful for understanding the data but cannot isolate the relationship in (6).

Figure 1. Reported energy-management measures in the primary sample



Notes: Unweighted shares in the 810-establishment primary sample. Each establishment contributes a baseline and a follow-up response. Baseline refers to the 2019-labelled cohort; actual fieldwork and recall periods differ across countries. The markers describe reported recent action, not a stock of installed capability.

The broader matched sample produces an apparently striking EU/non-EU contrast: the EU share rises from 35.1% to 49.1%, while the share in the five non-EU comparison economies declines from 36.8% to 28.3%. The unadjusted difference in changes is 22.5 percentage points, with a conventional country-clustered p-value of 0.092; adding baseline controls reduces it to 19.7 points, with $p = 0.122$. These are descriptive comparisons of

different groups, not estimates of the gas-shock mechanism. EU membership is not substituted for the predetermined exposure index, and this contrast is not used as the paper's causal finding.

6.2. Exposure and reported energy management

Table 3 reports the main exposure specifications. The unweighted coefficient is -0.0971, with a conventional 95% confidence interval of [-0.2614, 0.0672]. The independent wild-cluster p-value is 0.2344. Thus, the estimated gradient is negative but statistically inconclusive. The results do not provide affirmative evidence that establishments in more exposed country-industry combinations increased their probability of reporting energy-management action more than otherwise comparable establishments.

Table 3. Exposure and changes in reported energy management

	(1) First difference	(2) Baseline outcome	(3) Sampling weights
Exposure index	-0.0971	-0.0653	-0.0009
Country-clustered SE	(0.0712)	(0.0510)	(0.1095)
Conventional p-value	0.2099	0.2363	0.9939
Conventional 95% CI	[-0.2614, 0.0672]	[-0.1829, 0.0523]	[-0.2533, 0.2515]
Wild-cluster p-value	0.2344	Not computed	Not computed
Baseline controls	Yes	Yes	Yes
Country and industry effects	Yes	Yes	Yes
Baseline outcome control	No	Yes	No
Baseline sampling weights	No	No	Yes
Establishments	810	810	810
Country clusters	9	9	9
R-squared	0.089	0.542	0.233

Notes: All models include baseline log employment, age, exporter status, foreign ownership and a major/very severe financing obstacle, together with country and baseline-industry fixed effects. The dependent variable is the change in the binary reported-action indicator. Column (3) uses baseline median-assumption survey weights, not attrition-adjusted weights. Wild inference in column (1) is an independent supplementary calculation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, using conventional country-clustered inference.

The scale translation is useful because the raw exposure unit is constructed. One standard deviation corresponds to 0.4066 index units. Multiplying the primary coefficient and interval by this quantity yields a point estimate of -3.95 percentage points and bounds of -10.63 and 2.73 points. This interval should not be summarized as proof of no response: it includes substantial negative gradients and modest positive ones under the maintained model. Nor can it exclude positive effects on unmeasured adaptation outcomes or effects obscured by measurement error and selection.

Including the baseline management outcome reduces the coefficient in magnitude to -0.0653, with p = 0.236. Baseline sampling weights instead produce -0.0009, with p = 0.994. The weighted result reinforces the lack of affirmative evidence for a positive gradient, while weakening any argument that the negative unweighted point estimate is a stable population effect. The change in magnitude is itself informative about the sensitivity of the pooled descriptive estimand to how establishments are weighted.

The original analysis also labels separate manufacturing-only and EU-only estimates as sensitivities. Both restrictions reproduce the primary sample exactly because nonmanufacturing sectors and non-EU countries lack the required exposure inputs. Their coefficients are therefore identical to column (1). They are not counted as additional robustness evidence in Table 3. Likewise, the restricted-recall specification does not provide a further estimate: too few countries remain once the year-based timing restriction is imposed.

6.3. Employment and other business outcomes

The employment coefficient in Table 4 is -0.0951, with a country-clustered standard error of 0.0222. Using the primary sample's common exposure standard deviation as a scale benchmark, this corresponds to -0.0387 log points, or approximately a 3.8% lower ratio of follow-up to baseline employment. This is a conditional relative association, not an estimate that the typical firm shed 3.8% of its workforce, not an annual rate, and not a count of jobs destroyed by the crisis.

Table 4. Exposure and secondary business outcomes

Change in outcome	Coefficient	Cluster SE	Cluster p	Wild p	N
Log permanent full-time employment	-0.0951***	(0.0222)	0.0027	0.0039	849
Capacity utilization (percentage points)	-2.3095*	(1.2188)	0.0947	0.0508	668
Electricity spending/sales (percentage points)	0.9081	(0.9022)	0.3436	0.4727	634
Any power outage (0/1 change)	0.0432	(0.0647)	0.5225	0.6055	843
Generator ownership or sharing (0/1 change)	-0.0905*	(0.0427)	0.0670	0.2227	842
Exporter status (0/1 change)	-0.0421	(0.0396)	0.3191	0.3789	840

Notes: All models include baseline log employment, age, exporter status, foreign ownership and a major/very severe financing obstacle, together with country and baseline-industry fixed effects. Every row uses nine country clusters and outcome-specific complete cases. Estimates are per exposure-index unit, not standardized effects. Wild p-values are supplementary restricted bootstrap calculations; conventional intervals remain in the source results. * p < 0.10, ** p < 0.05, *** p < 0.01 based on the conventional p-value. No deflated-sales or productivity estimate is available.

Supplementary tests strengthen the statistical distinction between the employment signal and the other measured margins, while leaving the identification limitations unchanged. The employment wild-cluster p-value is 0.0039, compared with 0.0027 conventionally. Adjusting the seven-outcome family yields 0.0273. In the common 810-establishment sample, the coefficient is -0.0741 (standard error 0.0203); without baseline log employment it is -0.0924 (0.0294). Excluding Cyprus and Malta, the two zero-share countries, gives -0.1007 (0.0224). Appendix C reports these additional estimates separately from the original regression output.

None of these comparisons establishes a mechanism linking management to employment. In particular, the paper does not regress employment on an instrumented management response, and the baseline and follow-up outcomes need not be synchronized within the interval. The result is that employment changes exhibit a more clearly negative exposure gradient in these specifications. A claim that failure to adopt management practices caused labor reductions would require a different research design.

Capacity utilization has a coefficient of -2.31 percentage points per exposure unit. Its conventional p-value is 0.095 and its supplementary wild-cluster p-value is 0.051. Generator ownership or sharing also has a negative point estimate, but its conventional p-value of 0.067 becomes 0.223 in the wild-cluster calculation. Neither finding remains below 0.10 after the seven-outcome Holm adjustment. Electricity expenditure relative to sales, outage incidence, and exporter status are imprecisely estimated, with no conventional rejection at 10%.

The secondary variables do not close the accounting gap between energy use and productive adjustment. A higher electricity-spending share could reflect prices, quantities, or lower sales. Capacity utilization is a reported operating measure, not physical output. Generator ownership or sharing does not establish fuel switching or the emissions profile of replacement power. The absence of deflated sales and consistent physical energy quantities means the analysis cannot determine whether the employment association coincides with a productivity gain, reduced output, or a change in the composition of production.

6.4. Exploratory heterogeneity and country influence

Table 5 interacts exposure separately with baseline size, financing obstacles, and exporting. These characteristics were selected because they could relate to implementation capacity, access to resources, and exposure to external demand. They are proxies rather than measurements of the fixed costs or expectations in equation (1). Each interaction regression retains the corresponding main effects and the other baseline controls.

Table 5. Exploratory heterogeneity in energy-management changes

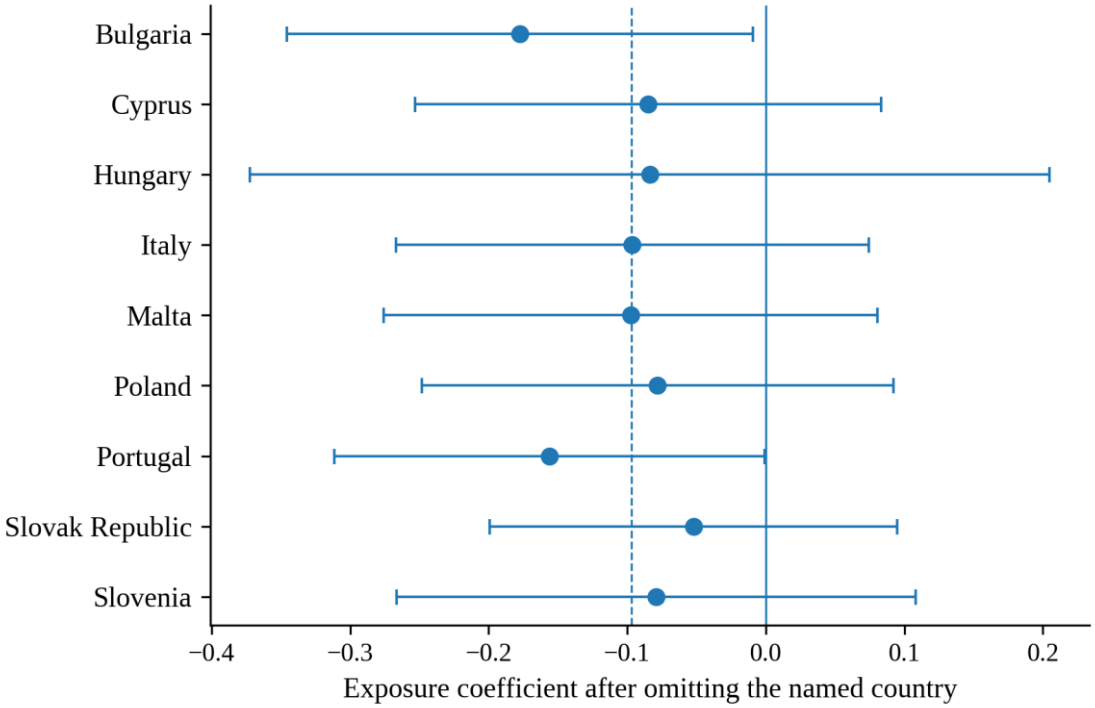
Exposure interaction	Interaction coefficient	Cluster SE	p-value	95% CI
Fewer than 100 baseline employees	0.0783	(0.1102)	0.4979	[-0.176, 0.332]
Major/severe financing obstacle	0.0984	(0.1081)	0.3891	[-0.151, 0.348]
Exporter at baseline	-0.2065	(0.1130)	0.1051	[-0.467, 0.054]

Notes: Each row is a separate specification with the interaction, its main effects, remaining baseline controls, and country and baseline-industry fixed effects. N = 810 and nine country clusters in all models. Coefficients are differences in exposure slopes, not subgroup treatment effects. No interaction is significant at 10%.

The estimated differential slope for establishments with fewer than 100 employees is 0.0783, with p = 0.498. The financing-obstacle interaction is 0.0984, with p = 0.389. The exporter interaction is -0.2065, but its p-value is 0.105 and its 95% interval includes zero. It is therefore not classified as statistically significant even at 10%. These results do not establish that a particular subgroup drove or avoided adaptation. The relatively small number of countries and the limited prevalence of reported financing obstacles constrain informative subgroup comparisons.

Figure 2 displays the original leave-one-country-out energy-management estimates. Every point estimate remains negative, ranging from -0.1775 when Bulgaria is omitted to -0.0522 when the Slovak Republic is omitted. Two exclusions produce conventional p-values below 0.05, but most do not. Selecting those exclusions would convert an influence diagnostic into specification search. The relevant finding is that no single exclusion turns the pattern into robust affirmative evidence of accelerated management action, while magnitude and conventional significance do vary.

Figure 2. Country-exclusion sensitivity of the energy-management coefficient



Notes: Original leave-one-country-out estimates, each using eight country clusters. Horizontal intervals are the reported conventional 95% confidence intervals. The dashed vertical line marks the full-sample coefficient; the solid line marks zero. Exclusions are influence diagnostics, not independent policy experiments.

Supplementary employment exclusions also leave the point estimates negative, between -0.1191 and -0.0823. The p-value rises to 0.052 when Bulgaria is omitted, illustrating that even the more precisely estimated employment relationship should not be described as uniformly significant at 5% under every exclusion. Country deletions test sensitivity to individual sample components; they do not address confounding shared by several countries or a misspecified technological benchmark.

7. Interpretation and implications for policy analysis

7.1. What the evidence says about adaptation

The combined evidence separates an observed organizational-action outcome from employment adjustment. In these data, greater predetermined exposure is not accompanied by a statistically established increase in reported energy-management measures, whereas changes in permanent full-time employment are negatively associated with exposure. This is compatible with the possibility that firms adjusted operating scale more readily than they introduced the specific organizational measures captured by the survey. It is not sufficient to

determine that operating contraction accounted for most energy savings or that technological adaptation was absent.

Several interpretations remain observationally consistent with the management result. Measures may have been adopted before the follow-up recall window, firms may have adjusted equipment or fuel use without describing the change as energy management, or common adaptation across Europe may not covary with recorded Russian import shares. Reporting differences can also blur the exposure gradient. Furthermore, establishments lost from the matched panel may have behaved differently from those retained. The paper therefore distinguishes the absence of affirmative evidence on this outcome from a substantive finding that firms did not adapt.

The employment association similarly has a narrower meaning than economic destruction. Permanent full-time employment can fall when firms reduce production, reorganize tasks, alter labor intensity, or substitute between employment arrangements. The WBES outcomes used here do not establish the contribution of each mechanism. This qualification is especially important because exposure is assigned at a broad baseline industry and country level rather than measured as the actual energy bill, contract renewal, or supply curtailment experienced by each establishment.

7.2. Energy-security shocks and climate-policy inference

The estimates inform the interpretation of energy shocks, but they do not rank climate or energy-security instruments. A temporary supply disturbance differs from a credible long-run price trajectory in its uncertainty, revenue consequences, and interaction with demand. An exposure gradient following the former is not an elasticity with respect to the latter. Converting the employment coefficient into a prediction for a carbon tax would require an independently measured price first stage and assumptions about comparability across shocks, technologies, and policy environments.

The distinction between action and performance also matters for policy evaluation. A measure of management adoption can document organizational activity, while metered energy use, emissions, output, and employment reveal different aspects of performance. A program could affect one without producing the expected movement in another. The present results cannot determine the relative effectiveness of information assistance, financing support, energy-price relief, or technology standards because those interventions are not

observed as distinct treatments. The conceptual framework indicates the channels such instruments could change, not which instrument performs best.

The lack of a measured cost-benefit comparison is equally consequential. Reductions in fossil-energy use may have environmental benefits, while disrupted production and labor adjustment can involve economic costs. Neither component is quantified here. There are no estimates of avoided emissions, investment costs, welfare, or displacement to production outside the observed economies. The empirical contribution is a more discriminating description of adjustment, rather than a welfare conclusion about whether an energy-security shock functions as climate policy.

7.3. Boundaries of the design and the next empirical test

Three extensions would directly address the principal identification gaps. First, linked administrative employment and sales records would help separate panel nonresponse from exit and improve the temporal resolution of adjustment. Second, establishment-level energy prices, contract dates, and metered quantities would allow the proxy exposure to be connected to an actual cost or supply first stage. Third, repeated, consistently worded measures of energy-management activity would support a pre-period comparison and distinguish recurring actions from an enduring stock of capability. These are data requirements implied by the present estimand, not results already established by the survey.

A revised technological benchmark would also need explicit validation. It should use a common fuel or first-use accounting boundary in the numerator and denominator and a documented crosswalk to the relevant industrial classification. Replacing the present proxy with such a measure would constitute a new exposure specification and require reestimating the reported models. Relabelling the current series as an exact gas intensity would not solve the underlying measurement issue.

Until those extensions are implemented, the defensible scope is the conditional relationship observed among reinterviewed establishments. The nine-country bootstrap makes the inferential treatment more transparent, but cannot supply missing counterfactual trends, correct nonrandom retention, or remove cross-country industrial dependence. The manuscript accordingly treats its strongest statistical result as evidence warranting a more granular employment investigation, not as a completed causal evaluation of the crisis.

8. Conclusion

This paper examines organizational and employment adjustment in a linked European manufacturing panel spanning the period before and after the gas crisis. It combines recorded Russian natural-gas import shares in 2020 with a predetermined U.S. gas-weighted energy benchmark and estimates first-difference exposure gradients with country and baseline-industry fixed effects. The primary management sample comprises 810 establishments across nine EU countries; the employment sample comprises 849.

Greater exposure is not associated with a statistically established increase in reported energy-management action. The primary coefficient is negative and imprecise, and baseline weighting moves it close to zero. Employment changes exhibit a negative association that survives supplementary common-sample, country-exclusion, and restricted wild-cluster checks. Other secondary outcomes do not show comparably stable statistical evidence after accounting for multiple outcome inspection.

The principal economic lesson is an empirical distinction: an energy-security episode can coincide with observable labor adjustment without establishing a corresponding organizational green transition. Here, that distinction is descriptive and conditional on the maintained design. Selective reobservation, the nine-country inference problem, the proxy character of exposure, nonidentical questionnaire wording, and incomplete timing information prevent stronger claims about causation or decarbonization. Future evidence linking actual energy costs, physical quantities, and consistent management measures to administrative firm outcomes would determine whether the observed employment gradient reflects temporary contraction, restructuring, or a more persistent change in production.

References

Adão, R., Kolesár, M., & Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, *134*(4), 1949–2010. https://doi.org/10.1093/qje/qjz025

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., & Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, *124*(1), 1–51. https://doi.org/10.1086/684581

Allcott, H., & Greenstone, M. (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives*, *26*(1), 3–28. https://doi.org/10.1257/jep.26.1.3

Amaglobeli, D., Guilhoto, J., Jahan, S., Khalid, S., Lam, W. R., Legoff, G. M., Meyer, B., Sheng, X. S., Smietanka, P., Waddell, S., & Weitz, D. (2024). *Firms' resilience to energy shocks and response to fiscal incentives: Assessing the impact of 2022 energy crisis* (IMF Working Paper 2024/027). International Monetary Fund. https://doi.org/10.5089/9798400263149.001

Auclert, A., Monnery, H., Rognlie, M., & Straub, L. (2023). *Managing an energy shock: Fiscal and monetary policy* (NBER Working Paper 31543). National Bureau of Economic Research. https://doi.org/10.3386/w31543

Bloom, N., Genakos, C., Martin, R., & Sadun, R. (2010). Modern management: Good for the environment or just hot air? *The Economic Journal*, *120*(544), 551–572. https://doi.org/10.1111/j.1468-0297.2010.02351.x

European Commission. (2022). *State of the energy union 2022: Snapshots per EU country*. Directorate-General for Energy. https://energy.ec.europa.eu/publications/state-energy-union-2022-snapshots-eu-country_en

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. https://www.jstor.org/stable/4615733

International Energy Agency. (2023a). *Gas market report Q1-2023: Including gas market highlights 2022*. OECD Publishing. https://doi.org/10.1787/cfa3e593-en

International Energy Agency. (2023b, March 14). *Europe's energy crisis: What factors drove the record fall in natural gas demand in 2022?* https://www.iea.org/commentaries/europes-energy-crisis-what-factors-drove-the-record-fall-in-natural-gas-demand-in-2022

MacKinnon, J. G., Nielsen, M. Ø., & Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, *232*(2), 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001

Martin, R., Muûls, M., de Preux, L. B., & Wagner, U. J. (2012). Anatomy of a paradox: Management practices, organizational structure and energy efficiency. *Journal of Environmental Economics and Management*, *63*(2), 208–223. https://doi.org/10.1016/j.jeem.2011.08.003

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, *92*(1), 160–180. https://doi.org/10.1257/000282802760015658

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, *19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Ryan, N. (2018). *Energy productivity and energy demand: Experimental evidence from Indian manufacturing plants* (NBER Working Paper 24619). National Bureau of Economic Research. https://doi.org/10.3386/w24619

U.S. Energy Information Administration. (2021a). *2018 Manufacturing Energy Consumption Survey: Table 1.2. First use of energy for all purposes (fuel and nonfuel), 2018*. https://www.eia.gov/consumption/manufacturing/data/2018/pdf/Table1_2.pdf

U.S. Energy Information Administration. (2021b). *2018 Manufacturing Energy Consumption Survey: Table 6.1. Consumption ratios of fuel, 2018*. https://www.eia.gov/consumption/manufacturing/data/2018/pdf/Table6_1.pdf

World Bank. (n.d.). *Enterprise Surveys methodology*. Retrieved September 17, 2026, from https://www.enterprisesurveys.org/en/methodology

World Bank. (2019). *Business Environment and Enterprise Performance Survey: Manufacturing module 2019*. https://documents1.worldbank.org/curated/en/695401607937965357/pdf/Business-Environment-and-Enterprise-Performance-Survey-Manufacturing-Module-2019.pdf

World Bank. (2023). *Enterprise Surveys and B-READY 2023 questionnaire*. https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/methodology/ES_B-READY%202023%20Questionnaire.pdf

Appendix A. Variable definitions and sample mapping

A.1. Definitions used in the analytical sample

Table A1. Analytical variable definitions

Variable or construct	Operational definition
Energy-management action	One for an affirmative recent-action response; zero for a negative response. Baseline BMGc23d is a listed measure; later ge8d is a standalone item with an explicit environmental purpose. Negative nonresponse/inapplicable codes are missing.
Change in energy-management action	Follow-up action indicator minus baseline indicator. Values -1, 0 and 1. This is not the change in the stock of installed management capability.
Employment change	Natural log of permanent full-time employment at the follow-up fiscal-year endpoint minus its baseline log. Nonpositive or invalid employment is excluded; the interval is not annualized.
Electricity spending share	100 times nominal electricity expenditure divided by nominal sales. Ratios above 100% are excluded by the supplied cleaning rule; no physical-efficiency interpretation is assigned.
Capacity, outages and generators	Capacity utilization is a reported percentage. Outage and generator-ownership/sharing variables are binary indicators; differences use valid answers at both endpoints.
Baseline controls	Log permanent full-time employment; interview-year minus start-year age; positive direct or indirect exports; foreign ownership >=10%; financing obstacle coded major or very severe.
Size interaction	One if baseline employment is below 100, zero otherwise. Includes eligible reinterviewed micro establishments and is not the EU legal SME definition.
Country component	Recorded Russian natural-gas imports divided by total natural-gas imports in 2020. Missing for the five non-EU comparison economies; those missing values are not zeros.
Sector component	2018 U.S. fuel-consumption intensity multiplied by the natural-gas share of first-use energy, with the aggregation in equation (4). A gas-weighted energy proxy, not exact gas use per dollar.
Combined exposure	Product of the country component and the baseline-ISIC sector component. Fixed at baseline country and industry throughout the panel.
Baseline weight	WBES median-assumption sampling weight from the baseline survey. Used in a sensitivity model only; no attrition adjustment is applied.

Notes: Definitions reproduce the executed analysis and distinguish measurement interpretation from the names of variables in the original data package.

The sample is formed using complete observations rather than imputing unavailable exposure or outcome information. Negative survey codes are retained in raw-source fields but excluded from the corresponding harmonized measures unless explicitly valid responses. The primary outcome requires a valid answer at both dates; secondary regressions use outcome-specific complete cases. Main tables report the resulting observation counts rather than assuming all outcomes use an identical sample.

A.2. Technological crosswalk and proxy values

Table A2. Baseline industry mapping and the U.S. energy proxy

ISIC	Industry description	NAICS mapping	Proxy value	Main N
15	Food products and beverages	311,3121	0.7710	212
16	Tobacco products	3122	0.1143	0
17	Textiles	313,314	0.5738	17
18	Wearing apparel; dressing/dyeing of fur	315	0.2500	55
19	Leather; luggage and footwear	316	0.2000	11
20	Wood and wood products except furniture	321	0.5959	18
21	Paper and paper products	322	2.8172	9
22	Publishing, printing and recorded media	323	0.3333	21
23	Coke, refined petroleum products and nuclear fuel	324	1.2963	4
24	Chemicals and chemical products	325	2.2644	17
25	Rubber and plastics products	326	0.3922	41
26	Other non-metallic mineral products	327	2.8546	35
27	Basic metals	331	2.8025	15
28	Fabricated metal products	332	0.3922	160
29	Machinery and equipment n.e.c.	333	0.1649	106
30	Office, accounting and computing machinery	334	0.1200	0
31	Electrical machinery and apparatus n.e.c.	335	0.2176	22
32	Radio, television and communication equipment	334	0.1200	8
33	Medical, precision and optical instruments; watches/clocks	334,339	0.1380	7
34	Motor vehicles, trailers and semi-trailers	336	0.1828	18
35	Other transport equipment	336	0.1828	4
36	Furniture; manufacturing n.e.c.	337,339	0.1673	30

Notes: ISIC Rev. 3.1 two-digit divisions; the NAICS labels reproduce the supplied mapping. Proxy values are unchanged from the estimation data and follow equations (3)-(4). Divisions 16 and 30 have zero primary-sample observations, leaving 20 occupied divisions. Source inputs: U.S. Energy Information Administration (2021a, 2021b).

The ISIC divisions in this table use the baseline Rev. 3.1 classification. Later surveys can use a different industrial classification, but exposure is not reassigned using a post-treatment classification. The crosswalk is an analytical approximation, not an assertion that broad U.S. NAICS categories and European ISIC establishments have identical technologies. In particular, shared NAICS mappings create repeated benchmark values and reduce effective technological granularity. Recycling and nonmanufacturing activities have no assigned benchmark in the headline design.

Appendix B. Sample selection, timing, and pre-period coverage

B.1. Broad panel coverage

Table B1. Baseline and matched samples across the 14 economies

Economy	Baseline N	Reobserved N	EM pairs	Reobserved (%)
Albania	377	152	152	40.3
Bosnia and Herzegovina	362	131	125	36.2
Bulgaria	772	218	198	28.2
Cyprus	240	80	76	33.3
Hungary	805	295	290	36.6
Italy	760	56	56	7.4
Kosovo	271	120	98	44.3
Malta	242	95	95	39.3
Montenegro	150	87	85	58.0
Poland	1369	185	154	13.5
Portugal	1062	308	302	29.0
Serbia	361	131	127	36.3
Slovak Republic	429	125	111	29.1
Slovenia	409	196	184	47.9
Total	7609	2179	2053	28.6

Notes: All retained industries. The selected later wave is the latest available post-crisis wave in each supplied panel. Counts are unweighted.

These counts refer to all retained sectors in the 14-country source panel, whereas the primary exposure sample is the nine-country manufacturing subset. Reobservation percentages divide selected post-wave matches by baseline cohort counts. Neither the numerator nor the denominator is a count from a complete business register, and missing follow-up records are not classified as exits.

B.2. Actual fieldwork and fiscal-year endpoints

Table B2. Interview and fiscal-year ranges in the primary sample

Country	Baseline interview	Follow-up interview	Baseline fiscal year	Follow-up fiscal year	N
Bulgaria	2019-2020	2023	2017-2019	2022	103
Cyprus	2018-2019	2024	2017-2018	2023	27
Hungary	2018-2020	2023-2024	2017-2019	2022	170
Italy	2019	2024	2017-2019	2022-2024	34
Malta	2019	2024-2025	2017-2018	2022-2024	31
Poland	2019	2024-2025	2018	2023-2024	98
Portugal	2018-2020	2023-2024	2017-2018	2022-2023	215
Slovak Republic	2019-2020	2023	2017-2019	2022	60
Slovenia	2018-2019	2024	2017-2018	2023	72

Notes: Ranges use observed fields in the exported first-difference data. The table reports endpoints, not the dates at which individual management measures were adopted.

The timing table is computed on the 810-establishment primary sample. A labelled survey year is not necessarily the interview year. The employment variables refer to the establishment's reported completed fiscal year, while the management item refers to the three years before interview. The available full-recall flag identifies only six Maltese and 51 Polish observations in the primary sample. Restricting to those observations leaves two countries and cannot sustain the main nine-country comparison.

Earlier employment pairs from 2013 and 2019 are available in the consolidated panel for 1,049 establishments across ten economies. Only Bulgaria, Hungary, Poland, the Slovak Republic, and Slovenia belong to the available gas-exposure country set. Their broad employment-pair counts are 91, 86, 187, 36, and 79, respectively, before additional manufacturing and complete-case restrictions. The executed workflow does not report a pre-period exposure regression. Neither these counts nor descriptive country mean changes demonstrate parallel trends for the primary management outcome.

Appendix C. Inference, replication, and supplementary estimates

C.1. Restricted wild-cluster calculations

Let X contain the full design matrix in (6), including the exposure column. Fit the model imposing the null $\beta = 0$ to obtain restricted fitted values $\hat{Y}^0$ and residuals $\hat{u}^0$. For sign vector $v = (v_1, \ldots, v_G)$, with each country weight equal to -1 or 1, construct the bootstrap outcome

$$Y_i^b(v) = \hat{Y}_i^0 + v_{c(i)}\hat{u}_i^0. \quad \text{(C1)}$$

Refit the unrestricted model to $Y^b(v)$ and compute the bootstrap t-statistic using the same country-clustered covariance form as in the original model. The CR1 finite-sample multiplier is $G/(G-1)$ times $(N-1)/(N-K)$, where K is the rank of the design matrix. The two-sided p-value is the fraction of the 512 statistics at least as large in absolute value as the original statistic. Enumeration is over bootstrap weights, not a permutation of countries' treatment assignments.

The supplementary implementation first reproduces all seven unweighted reported coefficients and standard errors independently from the exported first-difference dataset. Coefficients agree within 10^{-9} and the standard errors within 10^{-7}, the latter allowing for rounded Stata CSV output. The supplementary p-values are not represented as results from a native execution of boottest. They are calculated in the accompanying Python audit and preserve the original covariates, samples, and country clusters.

Table C1. Supplementary country-cluster and multiple-outcome inference

Outcome	N	Conventional p	Wild p	Holm-adjusted wild p
Energy management	810	0.2099	0.2344	1.0000
Log permanent full-time employment	849	0.0027	0.0039	0.0273
Capacity utilization (percentage points)	668	0.0947	0.0508	0.3047
Electricity spending/sales (percentage points)	634	0.3436	0.4727	1.0000
Any power outage (0/1 change)	843	0.5225	0.6055	1.0000
Generator ownership or sharing (0/1 change)	842	0.0670	0.2227	1.0000
Exporter status (0/1 change)	840	0.3191	0.3789	1.0000

Notes: All models use nine countries. Wild p-values enumerate 512 restricted Rademacher sign patterns. Holm adjustment treats these seven tests as one exploratory family. Values are not computed from independent new data, and inference remains conditional on the regression and dependence assumptions.

Holm-adjusted values use the seven unadjusted wild-cluster p-values in the table as one family. There is no assertion of strong finite-sample validity beyond the validity of those

underlying tests. In particular, small-cluster leverage, the assumed country-level dependence structure, and unobserved common industry shocks remain relevant. Confidence intervals in the main regression tables are the conventional cluster-t intervals, not bootstrap-inverted intervals.

C.2. Supplementary employment checks

Table C2. Supplementary employment specifications and country exclusions

Specification	Coefficient	Cluster SE	p-value	N	G
Reported primary employment	-0.0951***	(0.0222)	0.0027	849	9
Common energy-management sample	-0.0741***	(0.0203)	0.0065	810	9
Without baseline log employment	-0.0924**	(0.0294)	0.0138	849	9
Exclude zero Russian-share countries	-0.1007***	(0.0224)	0.0041	790	7
Leave out Bulgaria	-0.0830*	(0.0354)	0.0516	738	8
Leave out Cyprus	-0.0929***	(0.0198)	0.0022	821	8
Leave out Hungary	-0.1191***	(0.0229)	0.0013	678	8
Leave out Italy	-0.0916***	(0.0232)	0.0056	815	8
Leave out Malta	-0.1025***	(0.0243)	0.0039	818	8
Leave out Poland	-0.0912**	(0.0309)	0.0214	735	8
Leave out Portugal	-0.0823**	(0.0334)	0.0435	632	8
Leave out Slovak Republic	-0.0887**	(0.0277)	0.0151	781	8
Leave out Slovenia	-0.0868**	(0.0345)	0.0402	774	8

Notes: G denotes country clusters. Dependent variable: change in log permanent full-time employment. These independently computed sensitivity estimates are additional to the original Stata results. Country-clustered standard errors use the same CR1 correction. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; no bootstrap p-values are claimed for these additional specifications.

The additional employment checks are computed from the same first-difference data and do not alter the original reported regressions. The common-sample restriction requires valid management responses at both dates. The no-baseline-size specification omits baseline log employment while retaining the other controls. The zero-share exclusion removes Cyprus and Malta. Each country-exclusion model retains the same specification on the remaining countries. These are sensitivity calculations, not additional independently identified shocks.