Appendix to Gerlagh, Mathys and Michielsen, "Energy Abundance, Trade and Specialization" (ej371_02)

Appendix 1. Data Description

Table A. Countries in the sample.

Abbreviation	Country
AUS*	Australia
BEL	Belgium
CAN	Canada
DNK	Denmark
FIN	Finland
FRA	France
WGR	Germany, West
ITA	Italy
JPN	Japan
NLD	Netherlands
NOR	Norway
SWE	Sweden
GBR	United Kingdom
USA	United States

Notes: * for Australia, there is no data on trade

Table A. Sectors in the sample.

Abbreviation	Description	ISIC Rev. 2 code
AGR*	Agriculture	10
FOD	Food and Tobacco	31
TEX	Textiles and Leather	32
WOD	Wood and Wood Products	331, without furniture
CHE	Chemicals	351+35, includes non-energetic energy consumption, i.e. energy carriers as feedstock
NMM	Non-Metallic Minerals	36
PAP	Paper, Pulp and Printing	34
IAS	Iron and Steel	371
NFM	Non-Ferrous Metals	372
MAC	Machinery	381+382+383
MTR	Transport Equipment	384
CST*	Construction	50
SRV*	Services	61+62+63+72+81+83+90
TAS*	Transport	71

Notes: * trade data are missing for agriculture and non-tradable sectors

Table A3. Variable Description.

Variable	Description	Dimension	Data source
Capital	Net capital stock, 1990 prices, U.S.$	year-country-sector	ISDB-E
Labour	Total employment	year-country-sector	ISDB-E
Energy	Energy consumption in kt of oil equivalents	year-country-sector	ISDB-E
Exports	Exports of goods in U.S.$	year-country-sector	ISDB-E
Imports	Imports of goods in U.S.$	year-country-sector	ISDB-E
Value added	Gross value added, 1990 prices, U.S.$	year-country-sector	ISDB-E
Energy Price	Energy price per ktoe, 1990 prices, U.S.$	year-country-sector	ISDB-E
Energy production	Primary energy production in kt of oil equivalents	year-country	WBDI 2009
Energy consumption	Primary energy consumption in kt of oil equivalents	year-country	WBDI 2009
Income	GDP PPP constant international dollars per capita	year-country	WBDI 2009
Land	Area in sq. km	country	IDB 2009
Population	Mid-year population in thousands	year-country	IDB 2009
Savings	Domestic savings per GDP	year-country	WBDI 2009

Notes: ISDB-E: International Sectoral Database with Energy, WBDI: World Bank Development Indicators, IDB: International Database from U.S. Census bureau (www.census.gov/ipc/www/idb/region.php)

Country-sector energy data:

Energy consumption (ISDB-E): Total Final Energy Consumption in kilo tonnes of oil equivalent (Ktoe), a common unit to express total energy consumption from different energy carriers.

Country energy data:

Energy consumption (WBDI 2009): Energy use refers to use of primary energy before transformation to other end-use fuels, which is equal to indigenous production plus imports and stock changes, minus exports and fuels supplied to ships and aircraft engaged in international transport. Source: IEA.

Energy production (WBDI 2009): Energy production refers to forms of primary energy--petroleum (crude oil, natural gas liquids, and oil from nonconventional sources), natural gas, solid fuels (coal, lignite, and other derived fuels), and combustible renewables and waste--and primary electricity, all converted into oil equivalents. Source: IEA.

Appendix 2. Robustness checks

Table A4 contains the results for regression equation (1) when we use the log of value added rather than the log of employment as dependent variable. As noted and commented on in the main text, the interaction term remains positive and significant and the coefficients are larger than in the main specification.

Table A4. Coefficients on energy interaction term in regression eqn (1), dep. var: .

Lagged production x Intensity (cont)	0.162***			
Lagged production x Intensity (dummy)		0.148***		
Self sufficiency x Intensity (cont)			0.010***	
Self sufficiency x Intensity (dummy)				0.089***
R^2 adjusted	0.650	0.645	0.625	0.621
N	958	700	958	700
N clusters	166	120	166	120

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A5 presents the results for regression equation (2) when we scale net exports by total trade rather than value added. The coefficients on the interaction term are highly significant but lower than in the main specification. The reason is that, on average in our sample, the total volume of trade (exports plus imports) exceeds a sector's value added so that the variation in the dependent variable decreases when scaling net exports by total trade. The values for the explained variation (R^2 adjusted) increase slightly, indicating that the model works equally well for the two normalizations of net exports.

Table A5. Coefficients on energy interaction term in regression eqn (2), net exports scaled by total trade.

Lagged production x Intensity (cont)	0.0851***			
Lagged production x Intensity (dummy)		0.0772***		
Self sufficiency x Intensity (cont)			0.0886***	
Self sufficiency x Intensity (dummy)				0.0808***
R^2 adjusted	0.663	0.663	0.671	0.671

N	760	760	760	760
N clusters	130	130	130	130

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A6 contains the results when we do not weigh observations by sectoral trade exposure. The coefficients on the energy interaction term are slightly smaller in absolute value, but the main conclusions remain unaltered. The results suggest that energy abundance is slightly more important for employment in energy-intensive sectors in the large countries in our sample.

Table A6. Coefficients on energy interaction term in regression eqn (1); dependent variable log employment, observations not weighted by trade exposure.

Lagged production x Intensity (cont)	0.0770**			
Lagged production x Intensity (dummy)		0.0882**		
Self sufficiency x Intensity (cont)			0.0573**	
Self sufficiency x Intensity (dummy)				0.0686**
R^2 adjusted	0.465	0.482	0.454	0.469
N	955	697	955	697
N clusters	172	126	172	126

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Tables A7 and A8 show the results when we allow the coefficient on the energy interaction term to change over time. Although the coefficients become slightly smaller and less significant towards the end of the sample period, we cannot reject the null hypothesis that the effect is equal in all periods. The results indicate that our findings are consistent over all periods, separately.

Table A7. Coefficients on energy interaction term in regression eqn (1) over time.

Energy abundance	Lagged production	Lagged production	Self-sufficiency	Self-sufficiency
Energy intensity	Continuous	Dummy	Continuous	Dummy
1970-1974	0.121***	0.106***	0.0767***	0.0665***
1975-1979	0.0932***	0.0846***	0.0787***	0.0715***
1980-1984	0.0951***	0.0897***	0.0926**	0.0905***

1985-1989	0.101***	0.0983***	0.0894**	0.0940**
1990-1994	0.113**	0.113***	0.0623	0.0694*
1995-1997	0.0954*	0.121**	0.0649	0.0858*
R^2 adjusted	0.629	0.624	0.620	0.615
N	955	697	955	697
N clusters	172	126	172	126
p (no time variation)	0.941	0.904	0.571	0.380

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A8. Coefficients on energy interaction term in regression eqn (2) over time.

Energy abundance Energy intensity	Lagged production Continuous	Lagged production Dummy	Self-sufficiency Continuous	Self-sufficiency Dummy
1970-1974	0.178***	0.156***	0.132***	0.120***
1975-1979	0.150***	0.135***	0.151***	0.141***
1980-1984	0.147***	0.130***	0.118**	0.119**
1985-1989	0.172***	0.148***	0.155***	0.150***
1990-1994	0.203***	0.178***	0.139***	0.133***
1995-1997	0.174***	0.163***	0.119**	0.121**
R^2 adjusted	0.675	0.672	0.665	0.668
N	712	712	712	712
N clusters	124	124	124	124
p (no time variation)	0.495	0.773	0.159	0.249

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Appendix 3. Energy prices

In this appendix, we test whether energy abundance affects the location of energy-intensive industries through energy prices. Energy prices are likely to be endogenous to the location of energy-intensive industries: the presence of a large energy-intensive manufacturing sector drives up demand for energy and may thus result in higher energy prices. Conversely, energy-intensive industries might lobby for lower energy taxes, giving rise to lower energy prices. Regressing employment, value added or net exports on an energy price x intensity interaction term therefore leads to a biased coefficient. In order to alleviate this endogeneity, we instrument the energy price x intensity interaction term with an energy abundance x energy intensity interaction term using a two-stage GMM procedure. The equations we estimate are

(A)

, (A)

where we instrument with . We construct from country-sector specific energy prices by calculating a weighted average across sectors, with the sector weights given by total sample-wide energy expenditures in that sector as a fraction of sample-wide energy expenditures over all sectors combined. Table A9 and A10 show the second- and first-stage results when we use employment as the left hand variable.

Table A9. Second-stage coefficients on energy price interaction term in regression eqn (A1); dependent variable log employment.

	OLS	GMM-IV	OLS	GMM-IV
Prices x Intensity (cont)	-0.0309	-0.62**		
Prices x Intensity (dummy)			-0.0173	-0.70***
R2 adjusted	0.611	n.a.	0.602	n.a.
N	799	799	583	583
N clusters	172	172	126	126
Kleibergen-Paap statistic		5.07		6.51

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A10. First-stage coefficients of GMM-IV in Table A9 on energy interaction term in regression eqn (A1) with dependent variable log employment.

Lagged production x Intensity (cont)	-0.16**	
Lagged production x Intensity (dummy)		-0.17**
N	799	583
N clusters	172	126
R2 adjusted	0.52	0.47

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

From Table A10, we see that higher energy prices are associated with lower employment in energy-intensive industries. The GMM-IV coefficients in Table A9 are much larger in absolute value and more significant than the OLS estimates suggesting that a large energy-intensive manufacturing sector increases energy prices in a country. Another possible explanation is that the prices in our data are a noisy measure of the true prices faced by manufacturers, and that energy abundance is a better proxy for these true prices than our observed prices. The first-stage results in Table A10 indicate that energy abundance is associated with lower energy prices: a one standard deviation increase in lagged energy production leads to a decrease in energy prices of 0.16 standard deviations. The negative coefficients on the price interaction term in the second stage lend support to the carbon leakage hypothesis.

Tables A11-A14 show the results for value added or net exports as dependent variable. The picture that emerges is similar as in the main text: the coefficients on the energy price and abundance interaction terms are larger in absolute value than if we use employment as left-hand variable.

Table A11. Second-stage coefficients on energy price interaction term in regression eqn (A1); dependent variable log value added.

	OLS	GMM-IV	OLS	GMM-IV
Prices x Intensity (cont)	-0.0651*	-0.92**		
Prices x Intensity (dummy)			-0.0539	-1.25***
R2 adjusted	0.610	n.a.	0.604	n.a.
N	808	808	590	590
N clusters	166	166	120	120
Kleibergen-Paap statistic		4.05		4.94

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A12. First-stage coefficients of GMM-IV in Table A11 on energy interaction term in regression eqn (A1) with dependent variable log value added.

Lagged production x Intensity (cont)	-0.14**	
Lagged production x Intensity (dummy)		-0.14**
N	808	590
N clusters	166	120
R2 adjusted	0.48	0.44

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A13. Second-stage ccoefficients on energy price interaction term in regression eqn (A2); dependent variable net exports scaled by value added.

	OLS	GMM-IV	OLS	GMM-IV
Prices x Intensity (cont)	-0.0559	-1.10**		
Prices x Intensity (dummy)			-0.0440	-0.98***
R2 adjusted	0.642	n.a.	0.641	n.a.
N	598	598	598	598
N clusters	124	124	124	124
Kleibergen-Paap statistic		3.2		4.85

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A14. First-stage coefficients of GMM-IV in Table A13 on energy interaction term in regression eqn (A2) with dependent variable net exports scaled by value added.

Lagged production x Intensity (cont)	-0.13*	
Lagged production x Intensity (dummy)		-0.14**
N	598	598
N clusters	124	124
R2 adjusted	0.47	0.44

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Appendix 4. Data-driven energy abundance and intensity

In this Appendix, we extend the empirical method from Midelfart-Knarvik et al. (2000) and Romalis (2004) by showing that one does not need a country-measure for factor endowments or a sector-measure for energy intensity, but one can directly estimate them from country-sector input data. A fuller description of the methodology can be found in a working paper version of this article (Gerlagh and Mathys, 2011). In the monopolistic competition model of Romalis (2004), given the elasticity of substitution between varieties each firm will supply the same output level at the price

, (A)

where and are country- and sector-specific (inverse) productivity parameters, and the sector-specific factor shares of capital and energy and , and the factor prices of capital, energy and labour in country at time , respectively. The number of varieties in country for sector at time decreases in the price . Aggregate output is given by

,   (A)

where and are productivity parameters inversely related to and ; is the aggregate use of capital in country in sector at time , is energy input, and is labour input. Aggregate production costs are given by . Cost minimization yields the following first order condition in logs:

.   (A)

This equation can be directly estimated. Data on the left hand side (relative production inputs) are available, and the right hand side can be identified through sector and country-specific effects. The econometric model becomes

.   (A)

The relationships (1) and (2) can then be applied using the estimated and instead of and . If the coefficient on the interaction term is positive, countries that use more energy-intensive production techniques in all sectors have higher employment in energy-intensive sectors than countries that tend to use less energy-intensive production techniques. This procedure thus allows us to relate differences in the energy-intensity of production across all sectors to the sector structure in different countries.

From (5), the country effects specify the relative price of energy compared to labour in each country, so can be interpreted as an alternative measure of energy abundance that we can use as a sensitivity check. Because energy abundance may differ between countries because of endowments or due to possibly endogenous differences in regulations, we also perform a GMM-IV estimation in which we instrument with the weakly exogenous interaction term that we use in the main text. The first-stage coefficient on the instrument then indicates whether countries that have high lagged energy production also use more energy-intensive production techniques than countries with low energy production.

Table A15. Second-stage coefficients on interaction term in regression eqn (1); dependent variable log employment.

	OLS	GMM-IV 1	GMM-IV 2
	0.06	0.26***	0.35***
R2 adjusted	0.613	-0.55	0.37
N	955	955	697
N clusters	172	172	126
Kleibergen-Paap statistic		21.34	20.34

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. For GMM-IV 1, the instrument for is country lagged production times sector intensity (cont.); for 2 country lagged production times sector intensity (dummy). All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A16. First-stage coefficients of GMM-IV in Table A15 on energy interaction term in regression eqn (1) with dependent variable log employment.

Lagged production x Intensity (cont)	0.27**	
Lagged production x Intensity (dummy)		0.22**
N	955	697
N clusters	172	126
R2 adjusted	0.77	0.76

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A15 shows the effect of energy abundance on employment in energy-intensive industries using the estimated country endowments and sectoral intensities. The OLS coefficient is positive, but smaller in absolute value than in Table 2 in the main text, and not significant at the 10% level. When we instrument the estimated endowments with lagged production, the GMM-IV coefficients are larger and more significant however, as in Appendix 3, in which we instrumented energy prices with lagged production. This result is consistent with an explanation that we raised in Appendix 3, i.e. that *de facto* energy abundance in energy-rich countries is lower than suggested by lagged energy production, because demand from large energy-intensive manufacturing sectors drives up the energy price in these countries, partially offsetting the decrease in prices from high energy production.

The positive and significant coefficients on in the first stage indicate that countries with high lagged energy production employ more energy-intensive production techniques across all sectors, or alternatively that they specialize in energy-intensive subsectors within each sector. Thus, in addition to the 'composition effect' that we emphasized in the main text, we also find evidence that energy abundance also generates a 'technique effect'.

Tables A17-A20 show the results for value added or net exports as dependent variable. The results are broadly similar to the main text and Appendix 3: the second-stage coefficient on the interaction term is larger with value added as the dependent variable, though the OLS coefficient is very small and insignificant when we use net exports divided by value added as dependent variable.

Table A17. Second-stage coefficients on interaction term in regression eqn (1); dependent variable log value added.

	OLS	GMM-IV 1	GMM-IV 2
	0.091***	0.24	0.58***
R2 adjusted	0.617	-20.39	-4.24
N	955	958	697
N clusters	166	166	120
Kleibergen-Paap statistic		6.66	6.75

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. For GMM-IV 1, the instrument for is country lagged production times sector intensity (cont.); for 2 country lagged production times sector intensity (dummy). All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A18. First-stage coefficients of GMM-IV in Table A17 on energy interaction term in regression eqn (1) with dependent variable log value added.

Lagged production x Intensity (cont)	0.16***	
Lagged production x Intensity (dummy)		0.13**
N	958	700
N clusters	166	120
R2 adjusted	0.64	0.64

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A19. Second-stage coefficients on interaction term in regression eqn (2); dependent variable net exports scaled by value added.

	OLS	GMM-IV 1	GMM-IV 2
	0.011	0.26	0.45***
R2 adjusted	0.634	-1.61	-2.92
N	712	712	712
N clusters	124	124	124
Kleibergen-Paap statistic		6.45	7.38

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. For GMM-IV 1, the instrument for is country lagged production times sector intensity (cont.); for 2 country lagged production times sector intensity (dummy). All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.

Table A20. First-stage coefficients of GMM-IV in Table A19 on energy interaction term in regression eqn (2) with dependent variable net exports scaled by value added.

Lagged production x Intensity (cont)	0.17**	
Lagged production x Intensity (dummy)		0.14***
N	712	712
N clusters	124	124
R2 adjusted	0.64	0.64

Notes: all variables in logs, * significant at 10%, ** significant at 5%, *** significant at 1%. Error terms are clustered by country-sector pair. All regressions include country-time and sector-time fixed effects and a full set of sector dummies interacted with country savings, income, land area and population.