THE HETEROGENEOUS IMPACT OF COAL PRICES ON THE LOCATION OF CLEANER AND DIRTIER STEEL PLANTS

François COHEN, Giulia VALACCHI

The Heterogeneous Impact of Coal Prices on the Location of Cleaner and Dirtier Steel Plants

Francois Cohen[a] and Giulia Valacchi[b]

November 2020

Abstract:

Climate policy will predominantly affect industries that primarily rely on fossil fuels, such as steelmaking. Within these industries, exposure may be different by country according to the energy-intensity of national plants. We estimate the effect of coal prices on steel plant location worldwide and production preferences for BOF, a polluting technology, and EAF, a greener one. A 1% increase in national coal prices reduces BOF installed capacity by around 0.37%, while it has no statistically significant impact on EAF capacity. We simulate the implementation of a stringent European carbon market with no border adjustment and find a non-negligible shift in steel production outside Europe, with a concomitant impact on the technologies employed to produce steel. If applied worldwide, the same policy would primarily affect production in Asia, which relies on BOF and currently benefits from lower coal prices than those expected to emerge in the future.

Keywords: Steel industry; firm relocation; technological change; energy prices; carbon market; coal.

JEL Classification: O13, O33, Q41, Q42

a: Department of Economics, Chair of Energy Sustainability and Barcelona Institute of Economics, University of Barcelona, Barcelona, carrer de John Maynard Keynes, 1, 11, 08034 Barcelona, Spain. Email: francois.cohen@ub.edu

b: World Intellectual Property Organization, Geneva, Switzerland.

1. Introduction

The Paris Agreement (2015) has set the ambitious objective of limiting global warming below 2°C. It entrusts the Parties to determine their national contribution to this target according to domestic circumstances and capabilities. While the Nationally Determined Contributions (NDCs) of the Paris agreement allowed solving the deadlock of past failed negotiations, the aggregate efforts listed in all the NDCs are largely insufficient to achieve the below 2°C target. In that regard, only a small number of independent carbon schemes are currently operative (e.g. the EU, Japan, California and its partners in the US and Canada). In all these schemes, the price of carbon is far below its social cost as estimated by integrated impact assessment models.[1] One important reason why carbon prices are low on these markets comes from the risk that regulated industries lose competitiveness if other countries do not implement similar schemes. In energy-intensive sectors, largely exposed to international competition, unilaterally implementing a carbon tax or trading scheme may push industries to relocate elsewhere.

In this paper, we estimate the effect of changes in coal prices on steel plant location worldwide. Steel represents 27% of all greenhouse gas emissions (GHG) from industry (IEA, 2017b). We explicitly focus on the impact of coal prices on the steel industry because 75% of all its CO_2 emissions come from the burning of coking coal in Basic Oxygen Furnaces (BOF) (Columbia Climate Center, 2012). However, we can expect that coal prices have heterogeneous impacts across production processes because a less coal-intensive process, Electric Arc Furnaces (EAF), can also be used to produce steel. EAF is a recycling technology that cuts GHG emissions by 75% with

---

[1] Nordhaus (2017) estimates that the social cost of carbon is around $31/tCO_2 eq. for the current period. Yet, in 2015, the traded price of carbon ranged from $1 to $13/tCO_2 eq. worldwide, and 90% of carbon permits were traded at less than $10/tCO_2 eq. (WBG and ECOFYS, 2015). In the EU, the carbon price oscillated between $3 and $10/tCO_2 eq. between 2013 and 2015 (Marcu, 2016).

respect to BOF.[2]

Looking jointly at the effect of energy price shocks on plant location and production preferences constitutes the main contribution of this paper. The two are likely to interact. Even though changes of location may be encouraged by the availability of low coal prices in some countries, the choice to relocate will be ultimately dependent on the cost of staying, which depends on the availability of low-pollution technologies or the potential for their development. Following increases in coal prices, we find relocation effects for the dirtiest steel firms as well as increases in the diffusion of the low-polluting technology. However, because the low-polluting technology is a recycling technology that requires scrap, there are limits to its diffusion even under high coal prices.

This paper relies on steel plant data (1960-2014) collected by James King and merged with data on coal prices. We use a unit-level econometric model to correlate the size of national steelmaking industries to coal price shocks. Our setting circumvents several major identification issues. First, we account for the difference between current and expected coal prices by approximating coal price expectations with autoregressive integrated moving average models (ARIMA). We then rely on a pre-sample mean estimator (Blundell et al., 2002) to account for coal price endogeneity and the risk of weak instrumentation. Complementary robustness checks include tests for instrument exogeneity and several specification changes, for example in the definition of coal prices, or the use of an alternative estimator (system GMM). We also make sure that our results are not driven by the concomitant evolution of the prices of the other main steel production inputs: iron ore, electricity and scrap.

We find that an increase in coal prices at national level has a negative effect on the size of steel

---

[2] This process also relies indirectly on coal since it employs electricity that is generated from coal: globally, more than 70% of the GHG emissions from power generation are caused by coal (IEA, 2017a).

manufacturing. In our preferred specification, a 1% increase in coal prices reduces BOF production capacity by around 0.37%, while it has no impact on EAF capacity. As a result, both the size of production and its composition are influenced by coal price regimes.

We indirectly assess the effect of the introduction of ambitious climate policies on national steel industries by making the simplifying assumption that a carbon market is equivalent to a coal price increase. We simulate the impact of the implementation of two climate policies: a multilateral carbon market in the EU with a more stringent carbon price than today and no border adjustment; and a multilateral carbon market that would apply to all the countries that we cover (around 75% of the steel produced worldwide). In these carbon markets, we set the price of GHG emissions at $31/tCO$_2$ eq. This is the estimate of the current social cost of carbon in Nordhaus (2017). We find that redistributive effects across countries can be as large as redistributive effects across technologies. The portion of world capacity that is BOF would only decrease by 0.9 percentage points with a stringent European carbon market. Much of this BOF capacity would be replaced by EAF capacity. However, the potential for a transition from BOF to EAF is much greater: we simulate a 13.3 percentage point reduction of the global share of BOF production if the carbon market was implemented everywhere. However, this would also have an impact on the location of plants. We find that Asian countries would reduce their overall market share by around 19% if the carbon market was global because Asian firms are more coal intensive and have built their steel industry on relatively cheaper energy.

This paper complements a large body of economic literature that has looked at the effect of energy prices or environmental regulation on firm performance and location. Effects were found to be either positive or negative depending on the policy under scrutiny and the sector analyzed (Iraldo et al., 2011). Recent studies interested in industry, especially energy-intensive industry, have

shown that environmental regulation tends to decrease output and profits (Aldy and Pizer, 2015b; Ho et al., 2008; List et al., 2003; Greenstone, 2002)[3] and/or reduce exports and increase imports (Aldy and Pizer, 2015a; Levinson and Taylor, 2008; Ederington et al., 2005). Therefore, the location of plants should be impacted by environmental regulation and/or energy prices (Wagner and Timmins, 2009; Kellenberg, 2009; Kahn and Mansur, 2013). Some recent papers have studied the impact of the EU-ETS on firm relocation (Dechezlepretre et al., 2014, Borghesi et al., 2016, Koch and Basse Mama, 2016). Both Borghesi et al. (2016) and Koch and Basse Mama (2016) find evidence of relocation caused by the EU-ETS particularly for these sector exposed to international competition. On the other hand, Dechezlepretre et al. (2014) find no evidence of carbon leakages triggered by the introduction of the carbon scheme.[4]

Evidence that energy prices foster the adoption of cleaner technologies has been found in very diverse industry contexts (e.g. Cohen et al., 2017; Aghion et al., 2016; Dechezlepretre et al., 2011; Popp, 2006; Brunnermeier and Cohen, 2003; Popp, 2002; Gray and Shadbegian, 1998; Newell et al., 1999; Jaffe and Palmer, 1997; Lanjouw and Mody, 1996). Yet, none of the above-mentioned studies considered plant location and technological choice in the same framework.[5]

Even when alternative technologies exist, the risks of relocation, found in the case of BOF steel manufacturing, may apply to other highly energy-intensive industries (such as the paper or the chemical industry) that are also strongly exposed to international competition because of the

---

[3] Earlier studies attempting to measure the effect of environmental regulation on net exports, overall trade flows, job creation or plant-location decisions produced estimates that were either small or statistically insignificant (Jeppesen et al., 2002; Morgenstern et al., 2002; Jaffe et al., 1995).
[4] A limitation of their study is that they focus only on the initial period of activity of the EU-ETS, up to 2009, and they do not rule out the possibility of finding different results in case of an increase of the policy stringency by the EU policy makers.
[5] For the steel sector, Reppelin-Hill (1999) and Schleich (2007) have looked at the impact of energy prices on firm production choices. However, they do not properly address the endogeneity caused by the simultaneous determination of production choices and input prices. Mathiesen and Maestad (2004) aims to jointly account for the locational and technological choices of manufacturers in a static numerical partial equilibrium model of the world steel industry. These authors however rely on expert opinion to assess the magnitude of the elasticities relevant to their modelling exercise, in particular the elasticity of steel production technologies to changes in energy prices.

high tradability of commodities.[6] They may largely explain the reticence of EU countries to increase the stringency of the EU ETS, or to allow for exemptions in the steel sector due to the risk of carbon leakage. National industrial interests are one of the main reasons why multinational carbon markets are not being put forward. Since countries will be asymmetrically affected by it, systems with quota allocations that take into account the current distribution of firms across countries and their energy intensity may be necessary if a multilateral agreement on a carbon market is to be found. These findings suggest that enhancing international support for the implementation of NDCs, in accordance with the principle of differentiated responsibilities, is of utmost importance.

The rest of the paper is structured as follows. Section 2 presents the data while providing a brief overview of the steel industry. Section 3 presents our estimation method. Section 4 comments on the results and the main robustness checks performed. Section 5 presents our simulation exercise and section 6 concludes.

2. Data

2.1 James King data on the steel industry

The steel plant data has been gathered by James King and provides information on the location of steel plants in the world. For a few countries, the data is available since the beginning of the 20th century but it starts by around 1960 for most of them. The most disaggregated layer of observation in the database is the production unit: a steel plant is composed of several units, which may become operative or close down at different moments in time, even if they are on the same site. The data records the opening and closing year of these different units. These are permanent changes

---

[6] Some energy intensive industries produce goods that are more tradable (e.g. steel) than others (e.g. cement) because of the relative value of the commodity compared to transportation costs.

and we observe up to one opening data, and one closing date per unit.

Units may also use different production technologies. The dataset records whether a unit is EAF or BOF. BOF is a steel-making technology that came into wide adoption in the 1960s.[7] It produces steel with iron ore and coking coal.[8] 88% of CO_2 emissions associated with plants using BOF are due to the combustion of coking coal to obtain coke and then the mixing of iron ore with coke to obtain steel. The remaining CO_2 emissions indirectly come from electricity usage, usually generated with on-site coal-fired power generators (EPA, 2012; OECD, 2013; IEA, 2012). On the other hand, EAF is a recycling process that uses electricity from the grid to convert ferrous scrap into new steel (Giarratani et al., 2013).[9, 10] Emissions are mostly due to electricity usage in the EAF process (EPA, 2012).

As EAF uses recycled ferrous scrap and bypasses the coke production process, it is much less energy intensive. Switching from BOF to EAF is by far the option of greater impact on the emissions performance of steelmaking. EAF units emit, for the same amount of steel produced, only a quarter of the GHG emissions of their BOF counterparts (OECD, 2013 IEA, 2012). Other options have a much smaller emissions abatement potential.[11] However, since EAF resorts to scrap,

---

[7] BOF was invented in 1948 by Swiss engineer Robert Durrer and first commercialized in 1952–1953 by the Austrian steelmaking company VOEST and ÖAMG. However, the process only became widespread later in the 60s. In the U.S., the larger producer of steel, U.S. Steel, adopted this process in 1964 (Oster, 1982). Prior to the diffusion of BOF, the mostly-used steel-making process was Open Heart Furnace (OHF). It employed as main inputs raw materials such as iron ore, natural gas, oil or coal. It was a slow and inefficient procedure which, from the 1960s, has been completely replaced by BOF, which uses the same inputs, but exhibits big improvements in efficiency: approximately 1 BOF is required to replace 6 OHFs. Given this phase-out, we exclude OHF from the analysis.

[8] BOF can use scrap to produce steel, but up to a maximum of 25% of the amount of total inputs.

[9] The first EAF plant was established in the US in 1907 but, initially, the quality of the steel produced was lower than the one obtained through BOF and not enough scrap was around to make it cheaper to produce only through recycling. At the beginning of the 20th century, it was difficult to control the quality of the scrap, therefore, EAF steel was considered a byproduct. Innovations in the 60s and 70s allowed EAF to produce steel of good enough quality at a competitive price. In the U.S., Nucor introduced mini-mills in Darlington, CA, in 1969 and was followed by other manufacturers in the 70s (Schweikart and Doti, 2009).

[10] Steel scrap can have different sources: "home scrap" generated within the plant (nowadays it is not sufficient any more to produce steel due to the requirement of very high volumes of materials, therefore it needs to be integrated with the one purchased outside the firm); "new/prompt scrap" which is produced within the industrial activity of other firms (it is the same as home scrap, but it is not produced within the firm); "post-consumer scrap" which returns in the market after it ends its useful life (it could be very quick, as for cans, but it could take up to some years, as in the case of cars) (Yellishetty et al., 2011).

[11] Worrell et. al. (2001) discusses in details many alternative measures to improve the energy efficiency of steel manufacturing (e.g. preventative maintenance, cogeneration, heat recovery, etc). With 1994 data, assuming a payback period of 3 years and analyzing 47 specific energy efficiency technologies and measures, they found that energy efficiency could be improved cost-effectively in most iron and steel plants by 18%. In this paper, we have no data on the efficiency of active EAF and BOF units and we cannot evaluate the extent to which an increase in coal prices may lead to

it cannot fully substitute BOF. It is more frequently used in countries where scrap is sufficiently available to sustain production.

Units may substantially differ according to capacity. For example, BOF units tend to be much larger than EAF units. The James King data includes information on unit capacity. However, this information is only available for 2014. The information is missing for the units that closed down prior to 2014. For all the units for which we have capacity levels in 2014, we have assumed that historical capacity levels are equal to the 2014 capacity level. This is not a very strong assumption since we look at unit-specific capacity levels, not plant-specific levels. For the units that closed down before 2014, we have recovered unit capacity with a single imputation method.[12] More precisely, we ran a regression on all operative units in 2014. We took the logarithm of their capacity as the dependent variable. The explanatory variables used in the regression were technology fixed effects (BOF vs. EAF), technology by country fixed effects, and technology by year of installation fixed effects. This regression explained 66% of the variation in capacity across production units (R-squared of 0.66). We produced out-of-sample capacity estimates for the units that already closed down according to their technology, country and installation year (see Appendix A for more detail). In our baseline econometric specifications, 38% of observations have an installed capacity that is imputed. For the other 62%, the data includes information on capacity (in 2014).

Table 1 provides the descriptive statistics for installed capacity by technology, region and period after the imputation on the missing capacity data. We only display the data for the 22 countries that are part of the econometric analyses performed later. Other countries are not part of the analysis since

---

changes in the efficiency of installed EAF and BOF units. Still, our results are informative on the impact of coal prices on the environmental footprint of steel manufacturing since EAF production is far less carbon intensive.

[12] Our preference for a single imputation method over multiple imputations is because we use IV regressions later on. This process is not compatible with multiple imputations.

we do not have sufficient information on coal prices. In Online Appendix A, we report the capacity information for all countries, and deduct that the 22 countries that are part of our analysis represent around 75% of world production. We also report information on the number of units covered by our analysis, with a breakdown by technology and geography, before the imputation.

[TABLE 1 ABOUT HERE]

For 2000-2014, we estimate that about 70% of steel manufacturing capacity is BOF in our data. While total production capacity has not evolved dramatically in Europe and the US, it has significantly increased in Asia. For 2000-2014, most production capacity is in Asia according to our data. These capacity figures are indicative of production levels across countries and regions but are different from production data. Production is likely to be more responsive to immediate increases in coal prices. However, installed capacity should consistently reflect long-term trends in production levels resulting from changes in expected coal prices.

2.2 Coal price data

We use industrial coal price[13] data from the International Energy Agency, converted to constant 2010 USD,[14] except for China, for which we use the data from the Chinese Ministry of Coal, integrated with the Bohai-Rim Price Index. These sources report both coking coal and steam coal prices from 1978.

Steam coal, or thermal coal, is used primarily in electricity generation, while coking coal, or metallurgical coal, is used in steelmaking. In BOF processes, coking coal is converted into coke by eliminating virtually all impurities and leaving close to pure carbon. Around 600 kg of coke is

---

[13] We collected the total price charged to the industrial sector.

[14] Coal prices are initially extracted in national currency/tonne. To convert them in constant 2010 USD/tonne, we apply the formula: $P^{cons2010} = \frac{P_t^{LC}}{deflator_t^{GDP}} \frac{deflator_{2010}^{GDP}}{r_{2010}^x}$, where $P^{cons2010}$ is the price of coal expressed in constant 2010 USD/tonne, P_t^{LC} is the price of coal expressed in local currency/tonne at time t, $deflator_t^{GDP}$ is the GDP deflator for that country at time t, $deflator_{2010}^{GDP}$ is the GDP deflator of that country in 2010, and r_{2010}^x is the exchange rate between the local currency and USD in 2010. Prices are inclusive of taxes.

used to produce a tonne of steel, which means that around 770 kg of coking coal is used to produce a tonne of steel (WCA, 2017). The EAF processes do not involve mixing coke with iron-ore. However, they are reliant on electricity generated by coal-fired power plants. With EAF, around 150 kg of steam coal is used to produce a tonne of steel (WCA, 2017).

Figure 1 illustrates the evolution of average coking coal prices for a few countries (the United States, France, India and China) for 1978-2014. Coking coal prices tended to decline in the late 80s because of the concomitant oil price collapse (King and Tang, 1988). After 2000, they tended to increase because of the commodities super-cycle, mainly due to the rising demand from emerging markets such as the BRIC countries (Schwartz and Creswell, 2015). Coking coal prices have been consistently lower in Asia compared to Europe or North America.

[FIGURE 1 ABOUT HERE]

Even though EAF relies on steam coal and BOF on coking coal, we only include coking coal prices as an independent variable in our main specifications. We expect that shifts in coking coal price will capture the correlation between coking coal and steam coal prices, and therefore the impact of a change in steam coal prices on electricity prices. In Online Appendix D.3, we use steam coal prices and not coking coal in alternative regressions. Results with steam coal are very similar. This is not surprising since these prices strongly correlate. In addition, coal prices are instrumented in our estimations, which corrects for measurement error. In additional specifications (in section 4.2), we control for other input prices, among which electricity. As soon as electricity prices are controlled for, the effect of coking coal prices and electricity prices are separately estimated, making redundant the use of steam coal prices in the model.

2.3 Supplementary data

Some of our robustness checks make use of price data for the other inputs used in production: iron

ore for BOF, electricity and scrap for EAF.

For the electricity price we use the total price per MWh charged to the industry. The data has been obtained from multiple sources and merged in a unique database. Prices in national currency have been converted to constant 2010 USD. The sources for electricity prices are EUROSTAT statistics for the Eurozone countries[15], the Annual Report of the Chinese National Energy Administration for China, the Energy Price Report of South Africa for South Africa and the International Energy Agency for the remaining countries.

The data for iron ore and scrap comes from UN COMTRADE and has been calculated based on the import value and weight of goods in a given country.[16, 17] The data availability differ from one country to the other but goes back to the 1960s for a few countries. Measurement errors can be quite large because the COMTRADE nomenclature aggregates goods of different nature and because the price statistics is obtained by dividing value over weight. To reduce the risk of measurement errors, the data has been cleared. Only the observations with annual shipments of over 10,000,000 tonnes were considered to ensure higher homogeneity of goods. Furthermore, we excluded prices below the 5th and 95th percentiles. Online Appendix B provides summary statistics for the evolution of the price of electricity, iron ore and scrap over time. These price series have been deflated using the US consumer price index.[18]

3. Empirical methods

We aim to estimate if changes in coal prices lead to changes in the structure of steel manufacturing.

---

[15] EUROSTAT reports electricity prices charged to different categories of firms according to their annual electricity consumption. There are 9 categories before 2007, and 7 categories after 2007. Prices are averaged across all firms' categories.

[16] The choice of using import data was driven by data limitations: we could not find domestic price data for these two. In theory, shocks on import prices should align with shocks on domestic prices.

[17] For iron ore, we use code 281 ("iron ore and concentrates") of the standard international trade classification (SITC), and cross check the data with code 2601 of the harmonised system ("iron ores and concentrates, including roasted iron pyrites"). For scrap, we used code 282 ("iron and steel scrap") of the SITC.

[18] We use the US index since prices are in dollars. We prefer the consumer price index to the producer price index since the latter excludes imports, and therefore does not fully reflect changes in local prices.

We have good reasons to think that the price of coal will have an impact on the structure of steel manufacturing. First, coal is an essential component of steel manufacturing.[19] Second, there is evidence that manufacturers do not fully pass input and energy cost increases on to consumers through an equivalent increase in steel prices. Their margins are therefore affected and the level of coal prices should therefore determine investment decisions.[20]

We consider that, each year, manufacturers perform investment choices to open new production units where profitable and close old ones whenever expected profits are too low. Expected profits from these investment decisions are determined by the foreseeable cost of inputs – such as coal – and the price of outputs over the remaining lifetime of the plants. Therefore, the expected cost of inputs at time t – especially coal – have an impact on investment decisions at time t and should mechanically correlate with the overall size of steel manufacturing in a given region for this same period.

With this in mind, we develop an empirical strategy that allows estimating the impact of expected coking coal prices on the entry and exit of production units. Our empirical strategy is reduced-form: we do not directly model each step in the investment decisions and solely focus on the correlation between expected coal prices and the presence and size of production units. The main advantage of this approach is that it is less data demanding and does not impose restrictions on production functions to explain investment decisions. The main limitation is that we cannot model the effect of other parameters (e.g. the cost of switching from BOF to EAF) on the dependent

---

[19] Simple calculations show that the cost of metallurgical cost corresponds to a bit less than 25% of the price of crude steel in BOF processes: 780 kg of metallurgical coal is required to produce 1,000 kg of crude steel (Worldsteel association, 2019), and, in our data in 2014, the market price of coking coal is \$180 per tonne and the price of steel is \$587 per tonne: $(780/1000) * (180/587) \approx 23.9\%$.

[20] Ganapati, Shapiro and Walker (2016) find incomplete pass-through of energy costs in the US industry. Furthermore, they observe lower pass-through rates in industries with relatively low markups. Collard-Wexler and De Loecker (2015) explore markups in the U.S. steel industry and find that markups decreased significantly with the diffusion of EAF. The availability of two technologies to produce steel with different energy intensities constitute a strong limiting factor for the pass-through of energy costs in steelmaking.

variable, or the influence that these other parameters might have in explaining the correlation that we find between coal prices and steel plant location.

Our analysis is conducted at the level of production units. Our dependent variable is the share $N_{i,t}$ of operational capacity corresponding to unit i and year t, over the total worldwide capacity in year t: e.g. in 2005, unit i is operational and represents 1% of global capacity. Three factors have an influence on $N_{i,t}$. At the extensive margin, production units can either be operational or inactive. Inactive units correspond to those that have not entered the market yet, or that have closed down. Naturally, when units are inactive, $N_{i,t}$ is equal to 0. Furthermore, at the intensive margin, when an investor decides to invest in unit i, they can choose to install a unit with relatively small or large capacity. Finally, $N_{i,t}$ is also influenced by the total worldwide capacity in year t. This last phenomenon is however fully controlled for since we include time fixed effects in our econometric models.

The expected price of coking coal, denoted $p_{i,t}^*$, constitutes the independent variable of interest of our model. Therefore, we aim to estimate:

$$N_{i,t} = f(p_{i,t}^*)$$

Later on, we estimate this relationship for all production units altogether, and then separately according to the technology $s \in \{BOF; EAF\}$ used by the production units. This allows us to look at the impact of coal prices separately on BOF and EAF production units.

To estimate such a model, we are confronted with two difficulties. The first difficulty is that we do not observe the expected price of coal in country c at time t, but its actual realization. We denote $p_{i,t}$ the price of coal relevant to unit i at time t, according to its location (in country c). However, the right variable of interest is the expected price of coal because manufacturing companies are

likely not to base their investment decision on current inputs' prices. They will rather take in consideration their expected value over the lifetime of their investment. This is particularly true for a sector like the steel one because investments are costly and have a long lifetime.

The second difficulty is that expected coal prices should constitute an endogenous variable. A change in $N_{i,t}$ could come from a change in the demand for steel in country c at time t. This change in the demand for steel would affect the price of coal ($p_{i,t}$) because the steel industry is a large consumer of coal. Since $p_{i,t}$ also influences $N_{i,t}$, the two are simultaneously determined. On the other hand, expectations about future coal prices necessarily depend on the current price of coal. It follows that $N_{i,t}$ and $p_{i,t}^*$ are also simultaneously determined. Simultaneous determination will most likely lead to attenuated results. It could even produce inverted results because higher coal prices could be concomitant to higher domestic demand for steel. Another potential source of endogeneity comes from the risk of omitted variable bias. While our specification will control for time-invariant unobservables like in a fixed effect model, time-varying factors, such as changes in labour costs or water availability, might correlate with shifts in coal prices and confound our estimates.

The above-mentioned difficulties are likely to reinforce each other (and lead to attenuated and/or misleading estimates) if not carefully handled. Lack of information on expected prices creates measurement error, which exacerbates the endogeneity of expected coal prices. We handle these difficulties as follows.

3.1. Computing expected coal prices

To proxy expected coal prices, we consider that a perfectly rational agent forecasts future prices based on the information that s/he has. We assume that past prices constitute all the information easily available to economic agents, such that their expectations about future prices are a function

of past prices. In this framework, expectations can be recovered with a time-series model that produces a forecast of the coal price at time t+1, t+2, t+3, etc. with the coal price at time t, t-1, t-2, etc. We recursively apply autoregressive integrated moving-average models (ARIMA) on national annual time-series of real coal prices. The ARIMA model that we use to make the predictions is a first-order autoregressive model. This is because this model proved to be a good fit to the data.[21] This is a mean-reversing model: agents assume that price spikes or lows are unusual and expect that prices will align with longer-term trends. We use the following specification:

$$p_{i,t} = \mu_i + \gamma_i p_{i,t-1} + \varepsilon_{i,t}$$

In this model, the price of coal in each country i is regressed on its first lag. $\varepsilon_{i,t}$ is a term of error and μ_i and γ_i are parameters to be estimated. For each period t and country i, the ARIMA model is estimated with data from the previous 10 years. When we lack information on the previous 10 years, we estimate the model with all available years, provided that we have data for the previous 4 years at least. In each case, we then take the 10-year average of forecasted prices (for time $t+1$, $t+2$, $t+3$, etc.) as our value for the expected coal price, $p_{i,t}^*$. The ARIMA model is re-estimated for each time period t present in the sample, except for the first four years of our data (1978-1981), for which we did not have enough data to run the ARIMA model. We make out-of-sample predictions and, year after year, we allow steel companies to update their beliefs about future prices as soon as new information is available. Detailed statistics on the output of the ARIMA model and the predictions obtained are reported in Online Appendix C.

We performed several robustness checks relative to the choice of the coal price variable. They are

---

[21] In preliminary tests, we tried to use an AR2 process with the coal price data to derive expectations. However, the coefficient for the second lag in the ARIMA regressions was rarely statistically significant. There was therefore little justification for preferring an AR2 process. Nonetheless, in a complementary robustness check, we used an AR2 process instead of an AR1 process to derive coal price expectations and run our baseline models. Results were very similar. We chose not to report them for the sake of concision.

reported in Online Appendix D and also briefly discussed in the results section.

3.2. Endogeneity of expected coal prices

Our econometric strategy needs to deal with the endogeneity of the expected price of coal. In the literature, the most frequently suggested technique consists of using cost shifters as instruments, i.e. factors that are correlated with the cost of producing coal, and not with the demand for coal (e.g. Berry, 1994). However, the context of coal and steel production makes it hard to find strictly exogenous instruments: the two sectors are closely related. For example, steel production uses another output of the mining industry as an input, namely iron ore. Therefore, supply shocks on the extraction of coal may also affect the extraction of iron. If such cost shifters cannot be used as instruments, an alternative would consist in using shocks on the demand for coal that are not correlated with the demand from the steel industry as instruments. This approach is the one of Hausman et al. (1994), who instruments the price of a product in a given market with the price of this same product on other markets. Provided that demand shocks are not correlated across markets, this instrumentation strategy is valid. Yet, the steel industry is one of the main consumers of coal and markets are integrated across regions. When not used to make steel, coal is used to produce electricity. The assumption that demand shocks on the electricity market are not correlated with demand shocks on the steel market is not convincing: the two sectors are unlikely to be fully independent, with steel manufacturing using electricity, but above all largely dependent on economic growth in other sectors which consume electricity themselves, especially construction. We consider that most of the (demand and supply) shocks on coal prices are likely to be correlated with contemporaneous shocks on the steel industry. When contemporaneous correlations are strong, the usual approach consists in exploiting past information that is not correlated with contemporaneous shocks to produce valid instruments. Namely, it is possible to use the lags of the

endogenous variables as instruments, i.g. instrumenting $p_{i,t}^*$ with $p_{i,t-x}^*$ with x being an integer equal to 1 or above.

This choice, justified by the absence of better instruments, is not fully satisfactory at the outset because we cannot fully exclude that shocks in the deeper lags for expected coal prices may not correlate with the error term. Hereafter, we strongly reduce this risk by estimating dynamic models, i.e. expressing $N_{i,t}$ as a function of $N_{i,t-1}$ and $p_{i,t}^*$. This significantly reduces the risk that the exclusion restriction is not met, since any correlation between $N_{i,t}$ and $p_{i,t-x}^*$ should either go through $N_{i,t-1}$ or $p_{i,t}^*$. In complement, we naturally run standard over-identification tests to corroborate the fit of deeper lags for instrumentation. We also check that the inclusion of additional covariates has little influence on the estimated coefficient for the endogenous variable.

3.3. Available estimation methods

Traditional fixed effect models do not allow using pre-determined variables as instruments (e.g $p_{i,t-x}^*$ for $p_{i,t}^*$) because these rely on the assumption of strict exogeneity of the instruments.[22] Furthermore, they do not allow including $N_{i,t-1}$ as an independent variable in the model.[23] To circumvent these issues, researchers used to rely on models in first differences (Roodman, 2008). However, these models lack efficiency and, when time-persistent processes are studied, lags constitute weak instruments and the estimation is biased towards OLS results.[24]

There are two strategies to avoid weak instrumentation when instrumenting endogenous variables

---

[22] When predetermined-variables are used as instruments in a fixed effect model, they do not solve the endogeneity issue and results tend to be very similar to the results obtained without instrumentation.

[23] In Online Appendix E.2, we report the results that can be obtained with a fixed effect model. There is no treatment of endogeneity and we do not include $N_{i,t-1}$ in the model. As expected, results are inconsistent. They point to a positive correlation between coal prices and steel manufacturing. This positive correlation is not so surprising if we consider that coal price increases are likely to be concomitant to surges in the demand for steel (or economic growth in related sectors).

[24] In Online Appendix E.3, we provide an example of this. We use a model in first differences and instrument the log of the differenced expected coal prices, $\ln(p_{i,t}^*) - \ln(p_{i,t-1}^*)$, with their third lag, $\ln(p_{i,t-3}^*) - \ln(p_{i,t-4}^*)$. We show that instrumentation is weak and leads to inconsistent results. We also provide the (inconsistent) results of a model in first differences in which we assume that $\ln(p_{i,t}^*) - \ln(p_{i,t-1}^*)$ is exogenous.

with pre-determined regressors under time persistence. The first strategy consists in using pre-sample mean estimators, as in Blundell, Griffith and Windmeijer (2002). A specific interval of time $\{t_1; t_f\}$ is set to be the time of the analysis. In this interval, the set of information about dependent and independent variables is complete. There is, additionally, a pre-sample interval $\{t_i; t_0\}$ where only information on the dependent variable is available. A mean of the dependent variable is estimated over $\{t_i; t_0\}$ and included as a control variable in the estimation. Blundell, Griffith and Windmeijer (2002) show that pre-sample mean estimators are more efficient and significantly less biased than first difference estimators, even when time persistence is not extreme and for a small number of pre-sample observations. The intuition why pre-sample mean estimators are superior to first difference estimators is quite simple: they incorporate additional information into the model, namely the pre-sample mean of the dependent variable. The main limitation explaining why these estimates are barely used is that they can only be applied if pre-sample information on the dependent variable is available to the econometrician.[25]

We estimate the following log specification:

$$N_{i,t} = \alpha + \beta N_{i,t-1} + \gamma \ln(p_{i,t}^*) + \sigma \overline{N}_{i,t_i} + \theta_t + e_{i,t} \qquad (1)$$

Where $\overline{N}_{i,t_i}$ is the pre-sample mean of $N_{i,t}$, θ_t is a year fixed effect and $e_{i,t}$ is the error term. α, β, γ and σ are parameters to be estimated. Therefore, the pre-sample mean estimator is, in spirit, equivalent to a dynamic panel data model, in which we proxy the fixed effect with the pre-sample mean. As such, the model can globally be interpreted as a dynamic fixed effect model with year fixed effects. The long-term impact of the log. expected coal prices on $N_{i,t}$ is given by $\gamma/(1-\beta)$. We have data on $N_{i,t}$ since the 1960s, which is ideal considering that BOF and EAF processes

---

[25] This also requires stationarity. We test for stationarity using several unit root tests. Results are provided in Appendix G. These tests indicate that the dependent variable and coal prices seem stationary.

started to be widely used to produce steel from 1960 onwards. Since our data on $p_{i,t}^*$ starts in 1982, we have 22 years of pre-sample information on $N_{i,t}$ which can be used to run pre-sample mean estimators in the fashion of Blundell, Griffith and Windmeijer (2002). However, we use the period from 1970 to 1981 to calculate the pre-sample mean, and not the full period (1960-1981). The reason is simple. Our data is on unit openings and closings since 1960. Therefore, the 1960 data does not correspond to the stock of units that were active in 1960, but only to the flow that opened this year. After 10 years of data, i.e. in 1970, we can compute a value of $N_{i,t}$ that account for the relative importance of unit i compared to all the units that became active over the past 10 years and are still active in 1970. The pre-sample mean between 1970 and 1981 is therefore more representative of the share of capacity of unit i over total capacity in the pre-sample period.[26]

Equation (1) can be estimated using two-stage least squares (2SLS). To instrument for $\ln(p_{i,t}^*)$, the lags of $\ln(p_{i,t}^*)$ are valid instruments provided that contemporaneous shocks on $p_{i,t}^*$ are not correlated with previous shocks. The latter means that shifts in expectation between time t and t-x arise from the inclusion of new information about $p_{i,t}^*$. This is the case if expectations are rationally formed, or if they follow a random walk. In our base specification, we use $\ln(p_{i,t-5}^*)$ to instrument for $\ln(p_{i,t}^*)$. Our instrumentation strategy assumes that past shocks on expected coal prices do not correlate with the error term ($e_{i,t}$), but still determine the value of expected coal prices at time t. In mathematical terms, our instrument is valid if $corr\big(\ln(p_{i,t-5}^*), e_{i,t}\big) = 0$. The pre-sample mean estimator tolerates that $v_{i,t}$ is correlated with $e_{i,t-x}$ with $x \geq 1$, whereas a fixed effect model would require $corr\big(\ln(p_{i,t-5}^*), e_{i,t-z}\big) = 0$ with $z \geq 0$. In theory, we could use any lag, e.g. $\ln(p_{i,t-1}^*)$, to

---

[26] In Online Appendix E.1, we modify the pre-sample period used for estimation to 1960-1981 instead of 1970-1981. Results are similar to our preferred specifications.

instrument for $p_{i,t}^*$. In practice, using a deeper lag such as $p_{i,t-5}^*$ reduces the risk that $corr(p_{i,t-5}^*, e_{i,t}) \neq 0$, which would violate the exclusion restriction.

In an alternative specification, we expand the number of lags used as instruments, and run the model with $\ln(p_{i,t-3}^*)$, $\ln(p_{i,t-4}^*)$ and $\ln(p_{i,t-5}^*)$ as instruments. This allows us to run an over-identification test and make sure that all lags provide the same results when used as instruments. The general idea of this process is that all lags are valid instruments in theory and should therefore convey similar results. If the 3rd lag were not an exogenous instrument but the 5th lag were, then both could be providing different estimates. The result of this test, presented in section 4, corroborates our assumption that there is no correlation between $p_{i,t-5}^*$ and $e_{i,t}$ violating the exclusion restriction.

In another alternative specification, we instrument our model with the first differences of the lags of expected coal prices – i.e. $\Delta ln(p_{i,t-x}^*)$ with $\Delta ln(p_{i,t-x}^*) = ln(p_{i,t-x}^*) - ln(p_{i,t-1-x}^*)$ and $x \in \{2; 3; 4\}$.[27] This strategy of using first differences is often employed to reduce the risk that the exclusion restriction is violated when using lags as instruments. The results obtained with this choice of instruments match those obtained when using lags in levels, corroborating the validity of our instrumentation strategy.

The other econometric model to avoid weak instrumentation with pre-determined regressors under time persistence is the system generalized method-of-moments (GMM) estimator (Arellano and Bover, 1995; and Blundell and Bond, 1998). This estimator complements an equation in first differences with an equation in levels. This strongly increases precision compared to a standard first difference estimator and allows using lags in levels as instruments, which considerably reduces

---

[27] With x equal to 2, 3 and 4, we use up to the 5th lag to compute the instruments. This allows us to have similar estimation samples across all specifications.

the risk of weak instrumentation. Compared to the pre-sample mean estimator, the main advantage of system GMM is that is does not assume that the pre-sample mean is a good representation of the fixed effect. However, system GMM relies on the non-trivial assumption that first differences of instrument variables are uncorrelated with the fixed effects.

Hereafter, we provide the results obtained with system GMM alongside the ones obtained with the pre-sample mean estimator. This ensures that our results are not dependent on using one estimator instead of the other.

4. Results

Table 2 presents the results of the estimation of equation 1. Column 1 reports results of the impact of coal prices on installed capacity for the two technologies jointly. We find that an increase in expected coking coal prices leads to a reduction in installed capacity. Column 2 displays our results for BOF capacity and column 3 for EAF capacity separately.[28] They show that the impact of coking coal on operational capacity exclusively comes from BOF units. The long-term multiplier of the effect in column 2 implies that a 1% increase in the expected price of coal in country i would cause a decrease in available BOF capacity by about 0.37% (95% CI is 0.01-0.73%).[29] No effect is found for EAF units.

[TABLE 2 ABOUT HERE]

The results displayed in Table 2 rely on the validity of our instrumentation strategy. In Table 3, columns 1 (for BOF) and 3 (for EAF), we jointly use the third to fifth lags of $p_{i,t}^*$ as instruments

---

[28] The number of countries slightly vary in the samples used for estimation. The Netherlands leave the sample for EAF because there is no opening or closing of EAF units in this country in our sample. When 5 lags are used as instruments (Table 3), the number of countries in the pooled sample reduces from 22 to 19 (Russia, Spain and Switzerland leave the sample) since we do not have enough information on coal prices to compute all the lags. Later on, we use closer lags, or assume that the coal prices are exogenous (e.g. in the Appendices), and these countries come back in the sample.

[29] In absolute terms, a 1 percent increase in expected coal prices lead to an estimated reduction in the dependent variable that is equal to 0.0002389 percentage points. The sample average for the dependent variable is 0.0006472 for the estimation sample of Table 2, column 2. The reduction is of 0.37% (0.0002389 / 100 / 0.0006472).

and run over-identification tests. Results are very close to the ones obtained in Table 2 and the three over-identification tests of columns 1–3 corroborate the validity of our instrumentation strategy. In addition, we use the first difference of the lags of expected coal prices in columns 2 (for BOF) and 4 (for EAF) as instruments, instead of the lags in levels. In these specifications, we use the limited information maximum likelihood estimator to reduce the risk of weak identification. Results for Table 3, columns 2 and 4, are also very similar to our baseline results, suggesting that the exclusion restriction holds.

[TABLE 3 ABOUT HERE]

Besides, the results of Table 2 are only valid if the pre-sample mean is a good representation of the fixed effects. In Appendix E.1, we change the period used to compute the pre-sample mean and find similar results. We also split the sample period into two in Appendix E.4 to check that results are relatively stable across time. In Table 4, we also estimate a dynamic panel data model using system GMM. Results with system GMM are close to those of Tables 2 and 3, implying that our results do not rely on the choice of a pre-sample mean estimator. A difference, though, is that the impact of coal prices on BOF plants appears to be sharper.

[TABLE 4 ABOUT HERE]

Furthermore, the results of Tables 2–4 are robust to several specification changes. We check that these results do not rely on a specific functional form between coal prices and our dependent variable (see Online Appendices D.1). Results are not altered when we directly use expected coal prices instead of their logarithm in the model. When we add a quadratic term, results suggest that the marginal impact of an increase in coal prices is higher when coal prices are low, which is consistent with a logarithmic form. Also, our general results do not rely on the way expectations were computed. We obtain similar results, even though less precisely estimated, when we use

contemporaneous prices ($p_{i,t}$) instead of expected prices ($p_{i,t}^*$) as the main independent variable (see Online Appendix D.2). This implies that our results hold if we assume that manufacturers are myopic or think that coal prices follow a random walk.[30]

Finally, in Table 5, we assess if the impacts of Tables 2–4 come from changes at the extensive margin (i.e. the entry and exit of units). To do so, we reproduce the econometric estimation, but assume that any operational unit has a capacity normalized to 1. The new share $N_{i,t}$ is therefore equal to 0 for inactive units, or 1 for active units, divided by the number of operational units at time t. Results are provided with the pre-sample mean estimator (columns 1 and 2) and system GMM (columns 3 and 4), for BOF and EAF separately. Results in column 1 are statistically insignificant, while the other results point to a reduction in the number of all operative units, BOF and EAF, when coal prices increase. The difference found for BOF and EAF in Tables 2–4 therefore seem to come from changes at the intensive margin (i.e. changes in the capacity of operative units), and not from changes at the extensive margin (i.e. the number of units).[31]

[TABLE 5 ABOUT HERE]

4.2 Interpretation

We have found an impact of expected coking coal prices on BOF installations and no impact (pre-sample mean estimator) on EAF installations. These global estimates could in fact hide a large list of effects, both direct and indirect, of coal prices on installed capacity.

In the case of BOF units, the price of coking coal should influence installed capacity because BOF units directly use coking coal as an input. The model may also capture three additional effects.

---

[30] We also checked that excluding Chile and Australia, which are not in the same geographical areas as the other countries, had no influence on our general results. We do not report this robustness check for the sake of concision.

[31] As described before, installed capacity at unit level has been imputed based on the information available for 2014. Therefore, the imputation method creates measurement error in the dependent variable and may reduce efficiency. It has however no impact on model consistency under two assumptions: 1) unit capacities did not significantly differ between the units that closed down and the units that are still operative, conditional on technology, country and installation year; and 2) unit capacity did not go through substantial adjustments prior to 2014.

First, an increase in the price of coal may modify the relative competitiveness of BOF units compared to EAF units. Therefore, when coal prices increase, BOF units may be replaced by EAF units. Another effect comes from a change in the price of electricity due to a change in the price of coal. This effect is likely to be small but may increase production costs for BOF units. Finally, an increase in the price of coal could be correlated with the price of iron ore, since iron extraction is also reliant on energy.

In the case of EAF, our model could be capturing no effect at all, or a combination of effects cancelling out each other. There is no direct impact of coking coal prices on production costs since EAF units do not use coking coal. However, indirect effects could be occurring. On one side, EAF production could be increasing when coal prices are high because this technology becomes more competitive. On the other, several effects could reduce competitiveness gains, cancelling off any positive effect. For example, EAF uses electricity and, up to a certain extent, we could be capturing the effect that an increase in (steam) coal prices has on electricity prices. Above all, an important phenomenon is the interdependency between EAF and BOF production. EAF uses scrap as a main input to produce steel. In practice, the value of scrap should be correlated with the price of steel and therefore the price of BOF inputs. Besides, home market effects might also be important determinants. Both EAF and BOF technologies rely on a large downstream demand for raw steel. The downstream demand is likely to locate where steel manufacturing is to reduce transportation costs. Likewise, steel manufacturers have an interest to be geographically close to their clients. A downsizing of the BOF industry, which represents the largest share of the market, could significantly affect the smaller share of the market serviced by the EAF industry.

In Table 6, we aim to separately control for a series of effects that might explain the correlation between coal prices and BOF and EAF capacity levels. We add three additional control variables

for the other inputs likely to determine the level of installed BOF and EAF capacity: the expected

price of iron ore (used in BOF), electricity (mostly used in EAF) and scrap (mostly used in EAF).

[TABLE 6 ABOUT HERE]

Table 6, column 1, confirms that expected coking coal prices is the main input determining the

total capacity present in a given country. Column 2 shows that this effect is mostly driven by BOF

production, suggesting that all indirect effects may be rather small. Interestingly, when controlling

for other inputs, we find evidence of a small negative effect of coking coal prices on EAF. This

may stem from home market effects and the interdependency between BOF and EAF since EAF is

a recycling technology. The reader may also note that the stability of the results between Table 2

and Table 6 implies that our baseline specification is robust to the inclusion of additional covariates,

a feature that corroborates the validity of our instrumentation strategy.

5. Simulation of multilateral carbon markets

We now run a simulation where we quantify the effect of the introduction of multinational carbon

markets on the location of steel plants and the share of EAF units in a given country. Naturally,

this policy choice is illustrative since countries may opt for different types of policies to create a

carbon price. Cap and trade schemes and carbon taxes can be designed in very diverse ways. We

assume that if country i adopts the carbon market, it raises the price of carbon by an additional $

31 per tonne of CO_2 equivalent. This figure corresponds to the current social cost of carbon as

estimated in Nordhaus (2017). For simplicity, we assume that this increase in the price of carbon

would be additional to any existing policies, and that it would exclusively translate into an increase

in the price of coal. We convert the carbon tax of $ 31/tCO_2$ eq. into a coal price increase by

assuming that a tonne of coal emits $2.457/tCO_2$ eq.[32] We therefore raise the expected price of coal in regulated countries by about $ 76 per tonne.

We produce three scenarios with different combinations of countries implementing the scheme. In the business as usual (BAU) scenario, the policy is implemented nowhere. In the second scenario, the carbon market is adopted by the Member States of the European Union and Associated Countries (Norway and Switzerland in the simulation), with no border adjustment of steel prices with third countries. This scenario corresponds to the one where the steel industry would be strictly regulated under the EU Emissions Trading Scheme (ETS), with a much higher carbon price than the one we can be observe today. The last scenario assumes that all countries would implement the carbon market.

To assess policy impacts, we run a microsimulation based on the specifications displayed in columns 2 and 3 of Table 2: we therefore consider that the impact of increasing coking coal prices would be different for BOF and EAF units. We choose Table 2 since we know that the results of Table 2 are globally robust throughout Tables 3–5 and the appendices. Based on Table 2, we compute the share of BOF and EAF units that would have operated in country i over the sample period under the two policy scenarios, and compare this share to the one recorded in our data (which corresponds to the business-as-usual scenario).

Importantly, this microsimulation relies on equations that have been estimated in partial equilibrium. This type of analyses cannot account for general equilibrium adjustments. This is a strong limitation for two reasons. First, an increase in the price of coal by $ 76 per tonne would be clearly non-marginal: it corresponds to a 90% price increase (the average price of coking coal is $

---

[32] https://www.carbontrust.com/home/. Consulted June 2017.

84 in our sample).

Second, if all companies were subject to the same coal price increase around the world, investment decisions would be less sensitive to the price increase than if it were affecting only a small subset of countries. This is because if only a few companies face high coal prices, then they are at competitive disadvantage. However, if all companies faced the same increase in coal prices, then soaring coal prices would not create competitive imbalances across countries. In the case of a global carbon price or market, we would however suspect that three effects would still be at play and affect production capacity: 1) steel would be more expensive, so consumers would buy less steel; 2) other alternatives may become comparatively cheaper; 3) EAF steel would still gain competitiveness since it is less carbon-intensive. These three effects are likely to affect local production strongly because much of the steel consumed today is produced domestically. According to Deloitte (2018), the top 20 exporting countries of steel have exported 411 million tonnes in 2017, which only amounts to 24% of global output.

Overall, the fact that we rely on equations estimated in partial equilibrium implies that our results are likely to over-estimate total effects. The reader should therefore be careful. The output of our estimations are more informative about the direction of the change than they are about the exact market shares of steel manufacturing by location and technology under a high coal price regime.

On the other hand, in terms of technological change, our simulation only considers the switch from BOF to EAF. We would expect that carbon taxes lead to innovations within BOF and EAF technologies. The full technological response is therefore larger than the one that we can observe in this simulation.

With these limitations in mind, results are reported in Table 7. We are, above all, interested in shifts in the market share of BOF vs. EAF, since they would be very likely to happen even if all countries

would implement a carbon price. We are also interested in the likely location of plants in either North America, Europe or Asia. BOF represents 71.6% of installed capacity in the BAU. This would slightly decrease by 0.9 percentage points with a higher carbon market in Europe. However, if the carbon market is implemented everywhere, the reduction in the share of BOF capacity is strong, representing 13.3 percentage points (i.e. 18.6% in relative terms). While this value is large, it has been calculated for an even larger increase in coal prices (by 90%).

[TABLE 7 ABOUT HERE]

Estimated redistributive effects across countries are large as well. The share of BOF capacity in European countries would drop by 10.0% (from 16.0 to 14.4 percentage point) if they were the only ones to implement this carbon market. This could, however, be partly compensated by an increase in EAF production in Europe (increasing from 8.0 to 8.7 percentage points, i.e. by 8.3%). If the policy was implemented everywhere, we find that Asian countries would reduce their market share for both BOF and EAF production by 19.4% (minus 10.5 percentage points). The decrease would be due to a 35.6% reduction in BOF capacity, partially compensated by the increase in market share from EAF processes. This increase in Asian EAF capacity would be large in relative terms (60%), but much smaller in absolute terms (5.5 percentage points). The main reason why Asian countries would lose market share is because Asian firms are more coal intensive and they have built their steel industry on relatively cheaper energy.

We perform several robustness checks to assess the sensitiveness of these results. In Online Appendix F.1, we run the simulation with alternative carbon prices of $ 10 and $ 50, instead of $ 31. The magnitude of effects varies in proportion, but not our conclusions. In Online Appendix F.2, we run the simulation with a specification that uses the first differences of the lags of expected coal prices as instruments (Table 3, columns 2 and 4). We find similar results. We also run the

simulation with the GMM estimator (Table 5) in Online Appendix F.2. Results are very similar, but the drop in the production in Asia is more severe with this specification. In another robustness check, we use the expected price of coal in levels as the independent variable instead of its logarithm. We find simulation results that are similar to Table 7.

Finally, in the above simulation, we considered that marginal effects were homogeneous across regions. This assumption is relaxed in Online Appendix F.3, where we estimate region-specific response functions. With region-specific responses, we globally find the same trends as per Table 7 with the pre-sample mean estimator. We also find similar trends with a GMM estimator and region-specific effects. In particular, the very strong effect we found on Asian firms with the GMM estimator in Online Appendix F.2 is no longer present when we account for region-specific marginal effects.

Overall, this simulation is informative about the upcoming difficulties in reaching agreements on carbon markets across the globe. The steel industry would be sensitive to the implementation of a carbon market as a whole, since we predict a significant decrease in installed BOF capacity if the carbon price equaled the social cost of carbon of Nordhaus (2017). However, a higher carbon price would not affect countries equally. Our figures suggest that some European firms would relocate if the European steel industry was to comply with a more stringent carbon price with no border adjustments, e.g. in the framework of the EU ETS. Asian countries have benefited, so far, from a competitive advantage in cheap energy, which allows them to rely on polluting technologies to produce steel. They would lose part of this advantage if emissions were uniformly priced worldwide as they specialized more in BOF technologies. Finally, cross-country relocation effects appear to be at least as strong as within-country changes in steel-making technologies.

6. Conclusion

The Paris Agreement, in 2015, has shown the ambition of countries to undertake a global action against climate change. The steel sector is particularly sensitive to this topic as it is a major emitter of GHG emissions in the atmosphere. Given the availability of two substitutable steel-making technologies, the interest lies in the possibility, through environmental policies, to shift steel production from BOF to EAF, making it less carbon intensive.

This study aims at shedding light on the impact of an increase in coal prices on the location of steel plants and the technologies that the firms would choose to produce steel. In our baseline specification, we find that, on average, an increase of 1% in the expected price of coal reduces the share of BOF capacity by 0.37%, but has no impact on EAF capacity. When we simulate the effect of regional increases in coal prices, we find that European firms would lose competitiveness if they were to unilaterally set a binding carbon price on their firms, but this effect could be partially compensated by a shift from BOF to EAF production. Therefore, the risk of either relocation or asset stranding is real for EU firms if steel making processes were more heavily taxed. Another important finding is that a uniform increase of coal prices across the globe would also have an impact on the location of steel plants. Asian firms would be more severely affected because they are the ones that rely on BOF technologies the most.

The simulation on the European market shows that unilateral agreements may prove detrimental to national industries in industrialized countries, making the search for a multilateral agreement all the more necessary. In this regard, many Parties to the Paris Agreement have raised the concern that international support is currently insufficient and enhancing it will be necessary to increase the ambition of Nationally Determined Contributions (UNFCCC, 2016). A multilateral agreement could also have strong redistributive consequences for the steel sector. Our results suggest a

29

multilateral agreement on taxing energy use in the steel sector will be much easier to reach if measures were taken to guarantee that reducing GHG emissions in the industry would not strongly modify the location of production worldwide. Whereas these results mostly apply to the steel sector, they are likely to be valid for other energy-intensive sectors open to international competition, such as the paper industry. These redistributive effects would come on top of the reduction in demand that would stem from the increase in the prices of goods that require much energy to be produced.

Having said so, the reader should be careful that the general figures provided in the simulation are only indicative, since we are making extrapolations of effects for a non-marginal increase in coal prices, and do not account for macroeconomic adjustments and interactions at world level resulting from such an increase. As such, these results are however revelator of the sensitiveness of the steel industry to the implementation of ambitious climate policies.

References

Aghion, P., Dechezlepretre, A., Hemous, D., Martin, R. and VanReenen, J. (2016). Carbon taxes, path dependency, and directed technical change: evidence from the auto industry. *Journal of Political Economy.* 124(1), 1–51.

Aldy, Joseph E., and William A. Pizer (2015a). The competitiveness impacts of climate change mitigation policies. *Journal of the Association of Environmental and Resource Economists*, 2(4), 565-595.

Aldy, Joseph E., and William A. Pizer (2015b). The impacts of power sector carbon regulations on U.S. manufacturing employment. Harvard Kennedy School Working paper

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, *68*(1), 29-51.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics.* 25(2), 242–262.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115-144.

Blundell, R., Griffith, R. and Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics.* 108(1), 113–131.

Borghesi, S., Franco, C., and Marin, G., (2016). Outward foreign direct investments: pattern of Italian firms in the EU-ETS. SEEDS Working Paper.

Bosetti, V., Carraro, C., Duval, R., De Cian, E., Massetti, E. and Tavoni, M. (2014). The incentives to participate in and the stability of international climate coalitions: a game theoretic approach using the witch model. OECD Economics Department Working Papers (1177) .

Brunnermeier, S. and Cohen, M. (2003). Determinants of environmental innovation in us manufacturing industries. *Journal of Environmental Economics and Management.* 45(2), 278–293.

Chamberlain, G., 1992. Sequential moment restrictions in panel data. *Journal of Business & Economic Statistics* 10, 20-26.

Cohen, F., Glachant, M. and Soderberg, M. (2017). The impact of energy prices on product innovation: evidence from the UK refrigerator market. *Energy Economics*.

Columbia Climate Center, 2012. *Mitigating Iron and Steel Emissions*. The GHCN Factsheets.

Dechezlepretre, A., Gennaioli, C., Martin, R., and Muuls, M. (2014). Searching for carbon leaks in multinational companies. Grantham Research Institute on Climate Change and the Environment Working Paper (165).

Dechezlepretre, A., Glachant, M., Hascic, I., Johnstone, N. and Meniere, Y. (2011). Invention and transfer of climate change-mitigation technologies: a global analysis. *Review of Environmental Economics and Policy.* 5(1). 109–130.

Deloitte (2018) *Overview of the Steel and Iron Ore Market*, Deloitte CIS Research Centre, Moscow.

Ederington, J., A. Levinson, and J. Minier. 2005. Footloose and pollution-free. Review of Economics and Statistics 87 (1): 92 – 99.

EPA (2012). Available and emerging technologies for reducing greenhouse gas emissions from the iron and steel industry. Environmental Protection Agency.

Ganapati, S., Shapiro, J. S., & Walker, R. (2016). Energy prices, pass-through, and incidence in US manufacturing. *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-27*.

Giarratani, F., Hewings, G. J. and McCann, P. (2013), Handbook of Industry Studies and Economic Geography. Edward Elgar Publishing.

Greenstone, Michael. 2002. The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures. *Journal of Political Economy*. 110 (6): 1175-1219.

Hausman, J., Leonard, G. and Zona, J. D. (1994). Competitive Analysis with Differentiated Products. *Annals of Economics and Statistics.* (34), 143–157.

Ho, M. S., Morgenstern, R. D., & Shih, J. S. (2008). Impact of carbon price policies on US industry. *Resources for the Future Discussion Paper No. 08-37.*

IEA (2012). Profiles- co2 abatement in the iron and steel industry. International Energy Agency.

IEA (2016). Energy technology perspective 2016: Towards sustainable urban energy systems. International Energy Agency.

IEA (2017a). *CO2 emissions from fossil fuel combustion: highlights*.

IEA (2017b). *Tracking clean energy progress 2017.*

Iraldo, F., Testa, F., Melis, M., & Frey, M. (2011). A literature review on the links between environmental regulation and competitiveness. *Environmental Policy and Governance*, *21*(3), 210-222.

Jaffe, A. and Palmer, K. (1997). Environmental regulation and innovation: a panel data study. *The Review of Economics and Statistics*. 79(4), 610–619.

Jaffe, Adam B., Steven R. Peterson, Paul R. Portney, and Robert N. Stavins. 1995. Environmental regulation and the competitiveness of U.S. manufacturing: What does the evidence tell us? *Journal of Economic Literature.* 33 (1): 32–63.

Jeppesen, Tim, John A. List, and Henk Folmer. 2002. Environmental regulations and new plant location decisions: evidence from a meta-analysis. *Journal of Regional Science.* 42 (1): 19–49.

Kahn, Matthew E., and Erin T. Mansur. 2013. Do local energy prices and regulation affect the geographic concentration of employment? *Journal of Public Economics.* 101:105–14.

Kellenberg, D. K. (2009). An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of international economics*, *78*(2), 242-255.

King, R.E., Tang, H. (1988). International Macroeconomic Adjustment, 1987-1992 : a World Model Approach (The World Bank).

Koch, N. and Basse Mama, H. (2016). European climate policy and industrial relocation: evidence from German multinational firms. Available at SSRN: https://ssrn.com/abstract=2868283 or http://dx.doi.org/10.2139/ssrn.2868283

Lanjouw, J. and Mody, A. (1996). Innovation and the international diffusion of environmentally responsive technology. *Research Policy.* 25(4), 549–571.

Levinson, Arik, and M. Scott Taylor. (2008). Unmasking the pollution haven effect. *International Economic Review*. 49 (1): 223-54.

List, J. A., Millimet, D. L., Fredriksson, P. G., & McHone, W. W. (2003). Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator. *The Review of Economics and Statistics*, *85*(4), 944-952.

Marcu, A. (2016). Carbon market provisions in the Paris agreement (article 6). CEPS Special Report, Centre for European Policy Studies 128.

Mathiesen, L. and Maestad, O. (2004). Climate Policy and the Steel Industry: Achieving Global Emission Reductions by an Incomplete Climate Agreement. *The Energy Journal.* (4), 91–114.

Morgenstern, R. D., Pizer, W. A., & Shih, J. S. (2002). Jobs versus the environment: an industry-level perspective. *Journal of Environmental Economics and Management*, *43*(3), 412-436.

Newell, R., Jaffe, A. and Stavins, R. (1999). The induced innovation hypothesis and energy-saving technical change. *Quarterly Journal of Economics.* 114(3), 941–975.

Nordhaus, W. (2017). Evolution of modeling of the economics of global warming: changes in the dice model, 1992-2017. Cowles Foundation Discussion Paper No. 2084.

OECD (2013). Impacts of energy market developments on the steel industry.

Oster, S. (1982). The diffusion of innovation among steel firms: the basic oxygen furnace. *The Bell Journal of Economics*, 45-56.

Popp, D. (2002). Induced innovation and energy prices. *The American Economic Review.* 92(1), 160–180.

Popp, D. (2006). International innovation and diffusion of air pollution control technologies: the effects of nox and so2 regulations in the US, Japan and Germany. *Journal of Environmental Economics and Management.* 51(1), 46–71.

Reppelin-Hill, V. (1999). Trade and Environment: An Empirical Analysis of the Technology Effect in the Steel Industry. *Journal of Environmental Economics and Management.* 38(3), 283–301.

Roodman, D. (2008). Through the looking glass, and what ols found there: on growth, foreign aid, and reverse causality. Center for Global Development, Working Paper No. 137.

Schleich, J. (2007). Determinants of structural change and innovation in the German steel industry: an empirical investigation. *International Journal of Public Policy.* 2(1/2), 109–123.

Schwartz, N. D., Creswell, J. (23 October 2015). A Global Chill in Commodity Demand Hits America's Heartland. The New York Times.

Schweikart, L., & Doti, L. (2009). American Entrepreneur: A History of Business in the United States. Amacom.

Staiger, D., Stock, J. H., 1997. Instrumental variables regression with weak instruments. *Econometrica* 65, 557-586.

Stoft, S., Ockenfels, A. and Cramton, P. (2013). How to negotiate ambitious global emissions abatement. Working Paper .

UNFCCC (2016) *Aggregate effect of the intended nationally determined contributions: an update*. Report Number FCCC/CP/2016/2.

Wagner, U. J., & Timmins, C. D. (2009). Agglomeration effects in foreign direct investment and the pollution haven hypothesis. *Environmental and Resource Economics*, *43*(2), 231-256.#

WBG and ECOFYS (2015) *State and trends of carbon pricing*, 2015. World Bank Group.

WCA (2017). How is steel produced? World Coal Association. https://www.worldcoal.org/coal/uses-coal/how-steel-produced. Site consulted in June 2017.

Weitzman, M. L. (2014). Can negotiating a uniform carbon price help to internalize the global warming externality?. *Journal of the Association of Environmental and Resource Economists.* 1, 29–49.

Wooldridge, J. M., 1997. Multiplicative panel data models without the strict exogeneity

assumption. *Econometric Theory* 13, 667-678.

Worldsteel association (2019) *Factsheet: Steel and Raw materials*.

Worrell, E., Price, L., and Martin, N. (2001) Energy efficiency and carbon dioxide emissions reduction opportunities in the US iron and steel sector. Energy, 26(5), 513-536.

Yellishetty, M., Mudd, G. M., Ranjith, P. and Tharumarajah, A. (2011). Environmental life-cycle comparisons of steel production and recycling: sustanaibility issues, problems and prospects. *Environmental Science & Policy*. 4, 650–663.

Tables

Table 1: Descriptive statistics of James King data on steel capacity (in million tonnes) by location and main technology (1980-2014)

Period	1980-1989		1990-1999		2000-2014	
Technology	BOF	EAF	BOF	EAF	BOF	EAF
North America	85.7	51.3	81.1	70.8	70.2	88.0
Europe	219.4	58.5	218.2	84.1	186.7	122.6
Asia	260.2	22.9	306.4	58.0	594.9	113.7
Other	12.4	0.8	12.4	1.7	9.8	2.1
Total	577.7	133.4	618.1	214.6	861.5	326.5

Notes: The table reports the average installed capacity of steel-making production in the 22 countries covered in our statistical analysis. North America includes Canada and the United States. Europe includes Belgium, Czech Republic, Finland, France, Germany, Italy, The Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden and Switzerland. Asia includes China, India, Japan and Turkey. Other includes Australia and Chile. The information in this table is based on imputed capacity levels.

Table 2: Pre-sample mean estimation of the impact of expected coal prices on the share of installed capacity located in a given country

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	-0.00862***	-0.0105**	0.000968
coal price	(0.00197)	(0.00468)	(0.00193)
Lag dependent	0.961***	0.956***	0.963***
variable	(0.00318)	(0.00513)	(0.00179)
Pre-sample mean	-0.00200	-0.00327	-0.0186***
(1970-1981)	(0.00170)	(0.00309)	(0.00334)
Year fixed effects	Yes	Yes	Yes
Long-term multiplier for the	-0.2207***	-0.2389**	0.0258
effect of coal prices	(0.0527)	(0.1190)	(0.0508)
Weak identification test:			
Kleibergen-Paap rk Wald F	1180	1069	1373
Maximum IV size bias	<10%	<10%	<10%
Observations	41,394	14,506	26,888
Countries	19	19	18

Notes: The model is estimated with 2SLS. The logarithm of the expected price of coking coal is instrumented with its 5th lag. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

Table 3: Specifications using alternative instrumentation strategies

Dependent variable	Share of installed capacity worldwide (x 1,000)			
Technology	BOF		EAF	
Instruments	In levels [a]	In differences [b]	In levels [a]	In differences [b]
Column number	(1)	(2)	(3)	(4)
Log. expected coking	-0.0101**	-0.0112*	0.00105	0.00123
coal price	(0.00458)	(0.00604)	(0.00197)	(0.00229)
Lag dependent	0.956***	0.957***	0.963***	0.962***
variable	(0.00515)	(0.00506)	(0.00178)	(0.00189)
Pre-sample mean	-0.00336	-0.00296	-0.0187***	-0.0188***
(1970-1981)	(0.00307)	(0.00380)	(0.00334)	(0.00335)
Year fixed effects	Yes	Yes	Yes	Yes
Long-term multiplier for the	-0.2297**	-0.2562*	0.0279	0.0327
effect of coal prices	(0.1169)	(0.1460)	(0.0519)	(0.0599)
Weak identification test:				
Kleibergen-Paap rk Wald F	1876	24	1082	5
Maximum IV size bias	<10%	<10%	<10%	10-15%
p-value of Hansen J Statistic	0.13	0.14	0.20	0.18
Observations	14,504	14,504	26,884	26,880
Countries	19	19	18	18

Notes: In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

(a) The models in columns 1 and 3 are estimated with 2SLS. The logarithm of the expected price of coking coal is instrumented with its 3rd to 5th lags.

(b) The models in columns 2 and 4 are estimated using the limited-information maximum likelihood estimator. The logarithm of the expected price of coking coal is instrumented with the first difference of the 2nd, 3rd and 4th lags, i.e. $\Delta \ln(p_{i,t-x}^*)$ with $\Delta \ln(p_{i,t-x}^*) = \ln(p_{i,t-x}^*) - \ln(p_{i,t-1-x}^*)$ and $x \in \{2; 3; 4\}$. We choose the first differences of the 2nd to 4th lag so that the deepest lag we use to calculate the instruments corresponds to the 5th lag. Please note that only the instruments are in first difference, while the equations are in levels.

Table 4: Dynamic panel data estimation

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	-0.00927***	-0.0140***	-0.000304
coal price	(0.00167)	(0.00349)	(0.000950)
Lag dependent	0.969***	0.966***	0.968***
variable	(0.00812)	(0.00720)	(0.00234)
Unit fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Long-term multiplier for the	-0.2966**	-0.4152**	-0.0097
effect of coal prices	(0.1262)	(0.1853)	(0.0300)
Arrellano-Bond test for AR(2)	0.52	0.77	0.20
P-value of Hansen test	1.00	1.00	1.00
Degrees of freedom	233	223	233
Observations	52,080	18,286	33,794
Countries	22	21	21

Notes: The model is estimated with system GMM. The instruments are built with the 3rd to 5th lags of the lag dependent variable and the log. expected coal price. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

Table 5: Specifications using the number of units instead of installed capacity to construct the dependent variable

Dependent Variable: share of units	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Log. expected coking coal price	-0.000673	-0.00596**	-0.00000740***	-0.00000746***
	(0.00177)	(0.00271)	(0.00000194)	(0.00000191)
Lag dependent variable	0.960***	0.954***	0.975***	0.973***
	(0.00216)	(0.00324)	(0.00571)	(0.00518)
Pre-sample mean (1970-1981)	-0.0103***	-0.0167***		
	(0.00290)	(0.00171)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Long-term multiplier for the effect of coal prices	-0.017	-0.1289**	-0.3002**	-0.2786***
	(0.0447)	(0.0603)	(0.1395)	(0.0587)
p-value of Arrellano-Bond test for AR(2)			0.49	0.82
P-value of Hansen test			1.00	1.00
Degrees of freedom			181	193
Observations	14,506	26,888	18,286	33,794
Countries	19	18	21	21

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lag of the logarithm of expected coal price is used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lag of the logarithm of expected coal prices. In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels. We do not provide pooled results with BOF and EAF units altogether since the average capacity of BOF and EAF units is different.

Table 6: Estimation with additional inputs

Dependent Variable: share of installed capacity worldwide	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Log. expected coking coal price	-0.0241***	-0.00261	-0.0180***	-0.00373***
	(0.00733)	(0.00226)	(0.00502)	(0.00109)
Log. expected iron ore price	-0.00416	0.000467	0.000937	-0.00127
	(0.0228)	(0.00379)	(0.00935)	(0.00399)
Log. expected electricity price	-0.0245	0.00512	0.0153	0.00178
	(0.0204)	(0.0120)	(0.0142)	(0.00495)
Log. expected scrap price	0.0129	0.00155	0.0158	0.00330
	(0.0127)	(0.00409)	(0.0117)	(0.00329)
Lag dependent variable	0.952***	0.964***	0.950***	0.975***
	(0.00750)	(0.00165)	(0.00987)	(0.00626)
Pre-sample mean (1970-1981)	-0.00320	-0.0218***		
	(0.00407)	(0.00251)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Long-term multiplier for the effect of coal prices	-0.5022***	-0.0717	-0.3582***	-0.1473***
	(0.1602)	(0.0631)	(0.1536)	(0.0582)
Kleibergen-Paap rk Wald F statistic	35.8	55.5		
p-value of Arrellano-Bond test for AR(2)			0.96	0.08
P-value of Hansen test			1.00	1.00
Degrees of freedom			213	216
Observations	10,448	21,812	12,018	25,091
Countries	17	17	17	17

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 2nd lags of the logarithm of expected input prices are used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 2nd lags of the logarithm of expected input prices and the dependent variable. In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

Table 7: Simulation results

Indicator	World capacity that is BOF (%)			World capacity that is EAF (%)		
Scenario	BAU	Europe	All	BAU	Europe	All
World	71.6	70.7	58.3	28.4	29.3	41.7
North America	10.5	10.6	10.8	11.3	11.5	16.1
Europe	16.0	14.4	18.3	8.0	8.7	11.0
Asia	44.9	45.4	28.9	9.1	9.2	14.6

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia and Chile.

Figures

Figure 1: Evolution of coking coal prices in a few countries (1978-2014)
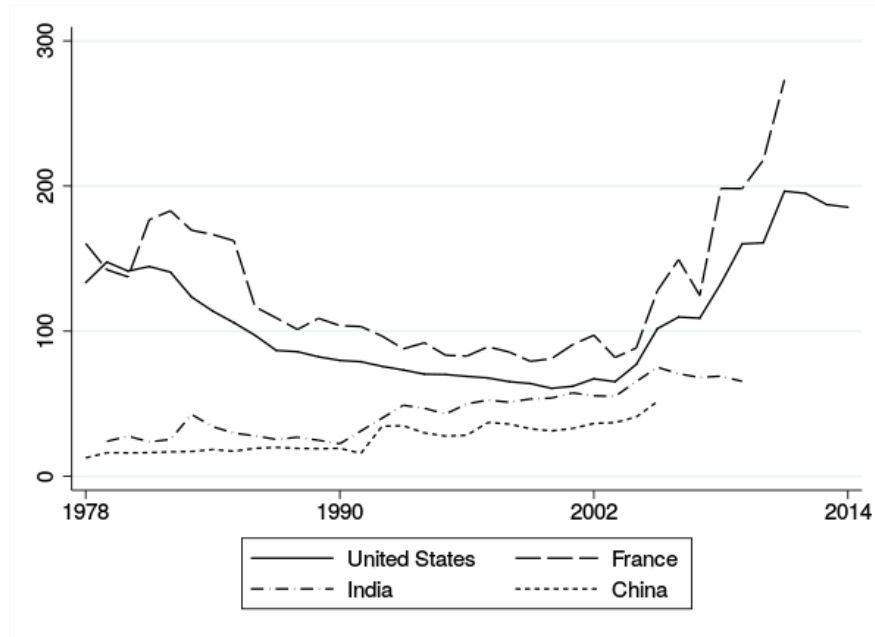
The Heterogeneous Impact of Coal Prices on the Location of Cleaner and Dirtier Steel Plants

Francois Cohen and Giulia Valacchi

Appendices

(Online Supplementary Material)

A. Imputation method for plant capacity

For all the units for which we have capacity levels in 2014, we have assumed that historical capacity levels are equal to the 2014 capacity level. This is not a very strong assumption since we look at unit-specific capacity levels, not plant-specific levels. For all the units that closed down prior to 2014, we have imputed a capacity value based on technology (BOF vs. EAF), country and year of installation. More precisely, we ran a regression on all operative units in 2014. We took the logarithm of their capacity as the dependent variable. The explanatory variables used in the regression were technology fixed effects (BOF vs. EAF), technology by country fixed effects, and technology by year of installation fixed effects. This regression explained 66% of the variation in capacity across production units (R-squared of 0.66). We produced out-of-sample capacity estimates for the units that already closed down according to their technology, country and installation year.

In the core of the text, we only provide the capacity data for the list of 22 countries for which we can conduct the empirical analysis. In Table A.1.1, we provide the aggregate capacity data for all countries. We observe that the 22 countries covered in our analysis represent 75% of world capacity.

Table A.1.1: Descriptive statistics of James King data on steel capacity (in million tonnes) by location and main technology (1980-2014)

Period	1980-1989		1990-1999		2000-2014	
Technology	BOF	EAF	BOF	EAF	BOF	EAF
North America	85.7	51.3	81.1	70.8	70.2	88.0
Europe	307.2	73.6	312.4	105.8	261.2	155.0
Asia	288.0	42.7	360.5	112.0	656.7	213.3
Other	67.1	27.5	74.5	46.9	77.3	63.7
Total	**748.1**	**195.0**	**828.6**	**335.5**	**1,065.3**	**520.1**

Notes: The table reports the average installed capacity of steel-making production in each region of the world over three different periods. Note that this information is based on imputed capacity levels, and covers all available observations in the James King data.

Instead of relying on installed capacity, we directly use the number of units in a robustness check displayed in Table 4. Table A.1.2 below provides information on the number of steel units recorded in the data for three periods (1980-1989, 1990-1999, 2000-2014), with a breakdown by technology and geographical location. We only display the data for the 22 countries that our empirical analysis covers.

Table A.1.2: Descriptive statistics of James King data on steel units by region and main technology used in production (1982-2014)

Period	1980-1989		1990-1999		2000-2014	
Technology	BOF	EAF	BOF	EAF	BOF	EAF
North America	70.3	233.0	64.4	209.5	53.8	182.6
Europe	132.9	236.2	126.0	227.3	108.0	233.7
Asia	199.2	156.8	232.6	217.7	351.1	295.3
Other	9.0	5.3	9.0	6.9	7.0	7.0
Total	**411.4**	**631.3**	**432.0**	**661.4**	**519.9**	**718.5**

Notes: The table reports the average number of active units of steel-making production in each region of the world over three different periods. We only report figures for the restricted list of 22 countries used in the regressions. North America includes Canada and the United States. Europe includes Belgium, Czech Republic, Finland, France, Germany, Italy, The Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden and Switzerland. Asia includes China, India, Japan and Turkey. Other includes Australia and Chile.

The majority of steel units in our sample are in Asia and there are about 50% more EAF units than BOF units. However, this information is misleading because BOF units have a much larger production capacity.

Table A.1.3 presents the share of total capacity (equal to 1 for each year) that was installed in four example countries (the US, France, India and China) according to the technology. There is a clear surge in Chinese production (both BOF and EAF) from the 90s, and a reduction in BOF production in the US.

Table A1.3: Share of total capacity by technology (BOF in left panel, EAF in right panel) for a few countries (1982-2014)
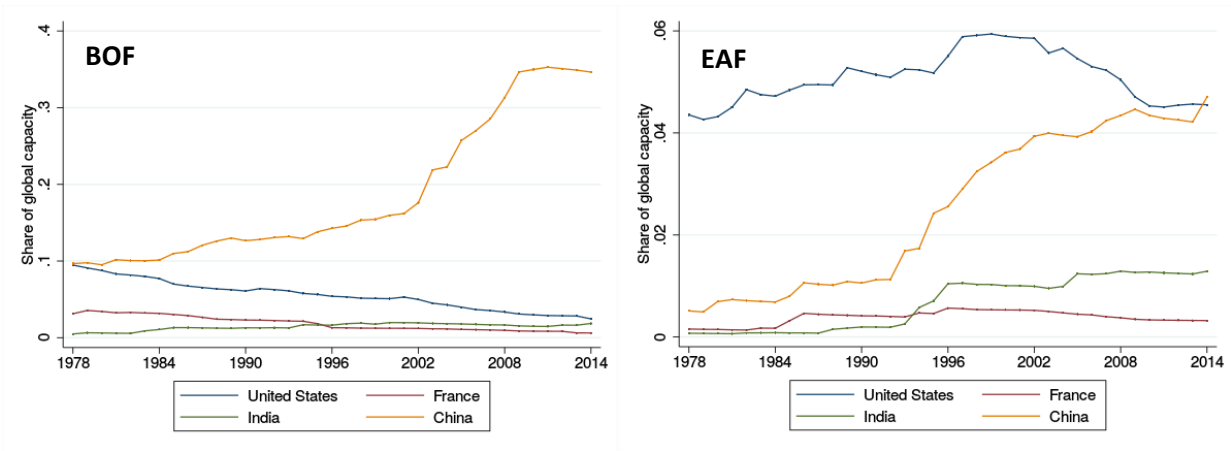
B. Price of other inputs

Figure B.1 below provide the average values of electricity, scrap and iron ore prices for the observations used in the estimation of the models with pre-sample mean of Table 6.

Figure B.1: Price of iron ore, scrap and electricity in the sample used to estimate the pre-sample mean estimator of Table 6 (1984-2014)
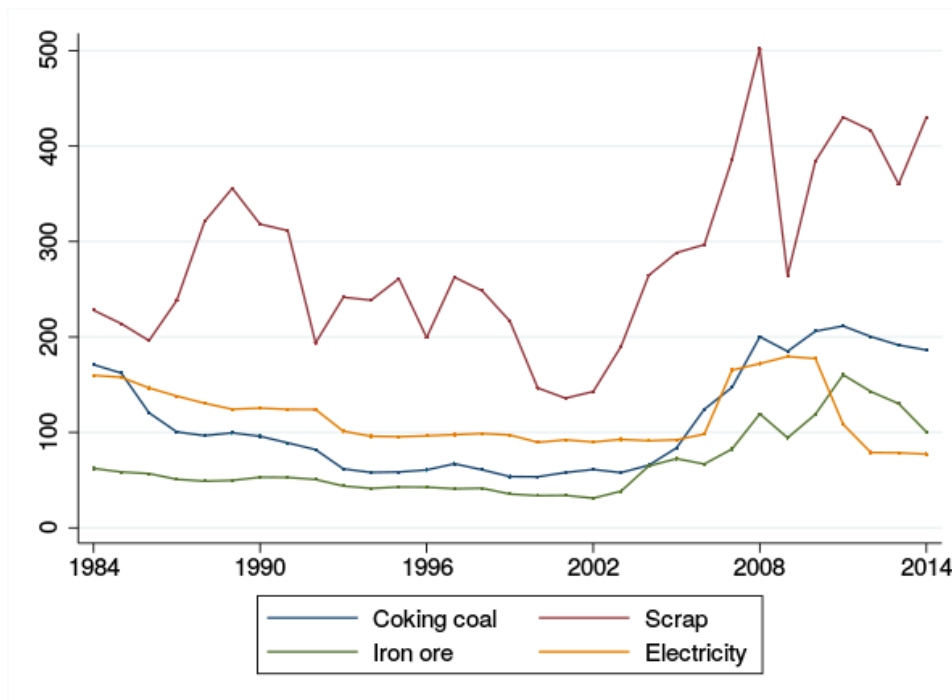
C. Output of ARIMA model to produce coal price expectations

Figure C.1 shows, for the US and China, the expectations on coking coal prices computed with the ARIMA model presented in section 3. Predictions are only displayed for specific years (1995, 2000, 2005, 2010 and 2015) even though they were computed for each year and country used to run our econometric model. The realized coal prices correspond to the black lines.

Figure C.1: Expected coking coal prices as predicted with the ARIMA model for the US (left panel) and China (right panel)
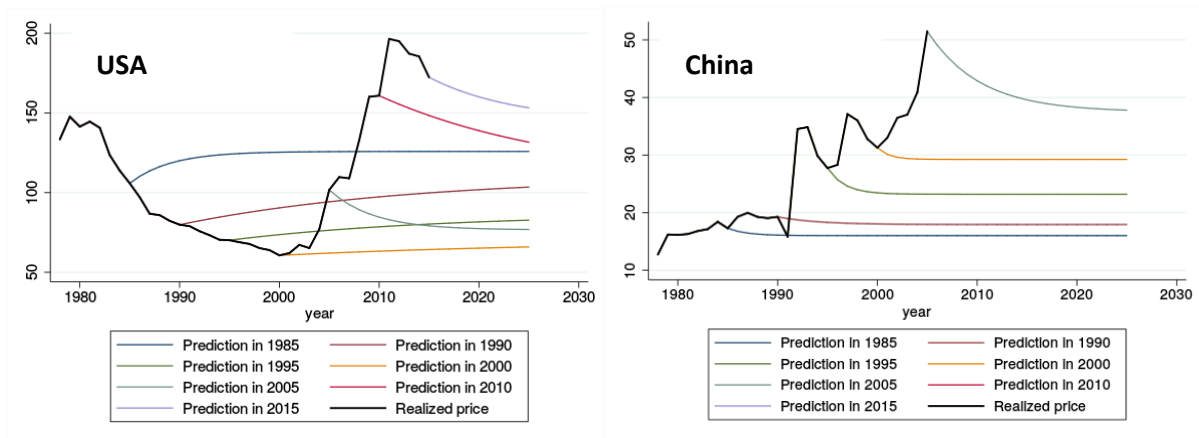
D. Alternative specifications of coal prices

1. Using coal prices in levels and not in logarithm

Table D.1.1 presents the estimation where the expected coal prices are included in levels instead of the logarithmic form.

Table D.1.1: Specifications using expected coal prices in levels

Share of installed capacity Worldwide (x 1,000)	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Expected coking coal price	-0.000129*	0.00003	-0.000189***	0.00002
	(0.00007)	(0.00004)	(0.00007)	(0.00002)
Lag dependent	0.956***	0.962***	0.961***	0.969***
variable	(0.00513)	(0.00195)	(0.00657)	(0.00232)
Pre-sample mean	-0.00417	-0.0188***		
(1970-1981)	(0.00296)	(0.00325)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
p-value of Arrellano-Bond test for AR(2)			0.80	0.20
P-value of Hansen test			1.00	1.00
Degrees of freedom			183	193
Observations	14,506	26,888	18,286	33,794
Countries	19	18	21	21

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lag of the logarithm of expected coal price is used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lag of the logarithm of expected coal prices. In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

In Table D.1.2, we have added a quadratic terms. Results suggest that plants are more sensitive to marginal increases in coal prices when they are low. This corroborates our preference for a model in logarithm.

Table D.1.2: Specifications using expected coal prices in level and their squared value

Share of installed capacity worldwide (x 1000)	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Expected coking coal price	-0.000393**	-0.00003	-0.0000003**	0.0000019
	(0.000162)	(0.00004)	(0.0000001)	(0.00003.85)
squared	0.0000014*	0.0000003	0.0000008*	0.00000005
	(0.0000007)	(0.0000002)	(0.000005)	(0.0000001)
Lag dependent variable	0.956***	0.962***	0.960***	0.969***
	(0.00514)	(0.00194)	(0.00716)	(0.00270)
Pre-sample mean (1970-1981)	-0.00298	-0.0183***		
	(0.00310)	(0.00289)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
p-value of Arrellano-Bond test for AR(2)			0.78	0.20
P-value of Hansen test			1.00	1.00
Degrees of freedom			295	305
Observations	14,506	26,888	18,286	33,794
Countries	19	18	21	21

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lags of the expected coal price in level and squared are used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lags of expected coal prices (in level and squared). In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

2. Using the contemporaneous price of coal

Table D.2 reports the estimation results obtained when using the current coal prices instead of the expected ones.

Table D.2: Specifications using contemporaneous coal prices

Share of installed capacity worldwide (x 1,000)	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Log. coal price	-0.00732	0.00195	-0.00739	0.000580
	(0.00000426)	(0.00000177)	(0.00636)	(0.00116)
Lag dependent	0.960***	0.964***	0.959***	0.968***
variable	(0.00455)	(0.00179)	(0.00680)	(0.00248)
Pre-sample mean	-0.00436*	-0.0169***		
(1970-1981)	(0.00247)	(0.00308)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
p-value of Arrellano-Bond test for AR(2)			0.93	0.12
P-value of Hansen test			1.00	1.00
Degrees of freedom			219	229
Observations	17,328	31,922	21,220	39,149
Countries	21	21	21	21

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lag of the logarithm of the coal price is used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lag of the logarithm of coal prices. In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

3. Using steam coal prices instead of coking coal prices

In Table D.3, we use expected steam coal prices instead of expected coking coal ones.

Table D.3: Specifications using expected stream coal prices

Share of installed capacity worldwide (x 1,000)	Pre-sample mean estimator		Dynamic panel data model	
	BOF	EAF	BOF	EAF
	(1)	(2)	(3)	(4)
Log. expected steam coal price	-0.0145	-0.00127	-0.0405***	-0.00561***
	(0.0124)	(0.00198)	(0.0104)	(0.00212)
Lag dependent	0.942***	0.962***	0.950***	0.973***
variable	(0.0144)	(0.00122)	(0.0126)	(0.00286)
Pre-sample mean	-0.00792	-0.0175***		
(1970-1981)	(0.00470)	(0.00286)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
p-value of Arrellano-Bond test for AR(2)			0.73	0.28
P-value of Hansen test			1.00	1.00
Degrees of freedom			193	193
Observations	13,761	27,139	17,607	33,930
Countries	20	21	22	23

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lag of the logarithm of the expected steam coal price is used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lag of the logarithm of expected steam coal prices. In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

E. Choice of estimation method

1. Using a different pre-sample period

Using a longer pre-sample period starting in 1960 has little influence on the results obtained (see Table E1 below).

Table E.1: Specifications with a pre-sample starting in 1960

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	-0.00807***	-0.00687*	0.000561
coal price	(0.00189)	(0.00355)	(0.00197)
Lag dependent	0.961***	0.955***	0.961***
variable	(0.00262)	(0.00429)	(0.00190)
Pre-sample mean	-0.00355**	-0.00554**	-0.00856***
(1970-1981)	(0.00126)	(0.00209)	(0.00175)
Year fixed effects	Yes	Yes	Yes
Long-term multiplier for the	-0.2092***	-0.1532*	0.0143
effect of coal prices	(0.0507)	(0.0865)	(0.0499)
Observations	41,394	14,506	26,888
Countries	19	19	18

Notes: The model is estimated with 2SLS. The logarithm of the expected price of coking coal is instrumented with its 5th lag. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

2. Results with a fixed effect model

In Table E.2, we present the results obtained when using a fixed-effect estimator. This model assumes that expected coking coal prices are fully exogenous. Results are biased and the model is inconsistent.

Table E.2: Specifications with a fixed effect model

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	0.205***	0.545***	0.0490***
coal price	(0.0265)	(0.0390)	(0.0112)
Unit fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	52080	18286	33794
Countries	22	21	21

Notes: The model is estimated with demeaning and the log expected coking coal prices is assumed to be exogenous. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

3. Results with a first difference estimator

Tables E.3.1 and E.3.2 feature first-difference estimators. Table E.3.1 uses a transformed model based on Chamberlain (1992) and Wooldridge (1997). In this model, instrumentation (with the third lag of the difference in log. expected coal prices) is weak. Weak instrumentation with lagged differences has been reported many times when using models in first differences with processes that are time-persistent. In this situation, we know that instrumentation is biased towards a model in first differences in which we would assume that the independent variables are exogenous (as shown in Staiger and Stock, 1997; and Blundell and Bond, 1998). In Table E.3.2, we report a model in first difference in which we assume that the expected price of coal is exogenous. All results lead to a positive association between coal prices and installed capacity, calling for a different modelling approach.

Table E.3.1: Specifications with a FD model based on Chamberlain (1992) and Wooldridge (1997)

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	0.304**	0.924***	0.0876**
coal price	(0.125)	(0.322)	(0.0410)
Year fixed effects	Yes	Yes	Yes
Kleibergen-Paap rk Wald F statistic	4.18	4.74	3.74
Maximal IV size bias	>25%	>25%	>25%
Observations	43,448	15,250	28,198
Countries	21	20	20

Notes: The model is estimated with 2SLS, using the third lagged difference of the log expected coking coal prices are an instrument. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

Table E.3.2: Specifications with a FD model with assumption of exogenous coal prices

	Share of installed capacity worldwide (x 1,000)		
	BOF & EAF	BOF	EAF
	(1)	(2)	(3)
Log. expected coking	0.0635**	0.184***	0.0164*
coal price	(0.0000224)	(0.0000523)	(0.00000794)
Year fixed effects	Yes	Yes	Yes
Observations	49919	17526	32393
Countries	22	21	21

Notes: The model is estimated with OLS, using the first difference of each variable. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

4. Splitting the sample into two periods

We split the sample into two and run the pre-sample mean estimator separately for 1982-1999 and 2000-2014. The pre-sample means are respectively obtained with the data from 1970-1981 and 1990-1999. Results indicate that the impact of coal prices on BOF production might have been stronger after 2000. However, results for the long-term multiplier are in fact not statistically different between columns 1 and 2.

Table E.4.1: Pre-sample mean estimation of the impact of expected coal prices on the share of installed capacity located in a given country

	Share of installed capacity worldwide (x 1,000)			
Technology	BOF	BOF	EAF	EAF
Sample period	1982-1999	2000-2014	1982-1999	2000-2014
(Pre-sample period)	(1970-1971)	(1990-1999)	(1970-1971)	(1990-1999)
Column	(1)	(2)	(3)	(4)
Log. expected coking	-0.00882**	-0.0527***	0.00170	-0.000298
coal price	(0.00381)	(0.0156)	(0.00149)	(0.00575)
Lag dependent	0.962***	0.944***	0.962***	0.969***
variable	(0.00788)	(0.0103)	(0.00255)	(0.000754)
Pre-sample mean	-0.00215	-0.00197	-0.0298***	-0.00664**
(1970-1981)	(0.00453)	(0.00653)	(0.00341)	(0.00243)
Year fixed effects	Yes	Yes	Yes	Yes
Long-term multiplier for the	-0.2294**	-0.9461**	0.0444	-0.0096
effect of coal prices	(0.0859)	(0.3872)	(0.0390)	(0.1844)
Weak identification test:				
Kleibergen-Paap rk Wald F	1,504	242	1,239	260
Maximum IV size bias	<10%	<10%	<10%	<10%
Observations	8,708	5,798	14,195	12693
Countries	16	13	15	13

Notes: The model is estimated with 2SLS. The logarithm of the expected price of coking coal is instrumented with its 5th lag. Cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

F. Simulation

1. Results using different carbon prices

Table F.1.1: Simulation results with a carbon price of $ 10 per tonne

Indicator	World capacity that is BOF (%)			World capacity that is EAF (%)		
Scenario	BAU	Europe	All	BAU	Europe	All
World	71.6	71.2	66.3	28.4	28.8	33.7
North America	10.5	10.5	10.8	11.3	11.4	13.2
Europe	16.0	15.4	17.1	8.0	8.2	9.2
Asia	44.9	45.1	38.2	9.1	9.1	11.3

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia (1987-1989) and Chile (2011-2014).

Table F.1.2: Simulation results with a carbon price of $ 50 per tonne

Indicator	World capacity that is BOF (%)			World capacity that is EAF (%)		
Scenario	BAU	Europe	All	BAU	Europe	All
World	71.6	70.3	52.4	28.4	29.7	47.6
North America	10.5	10.6	10.7	11.3	11.5	18.3
Europe	16.0	13.7	18.9	8.0	9.0	12.4
Asia	44.9	45.7	22.4	9.1	9.2	16.9

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia and Chile.

2. Results with alternative specifications

Table F.2.1: Simulation results using instruments in first difference and the pre-sample mean estimator (Table 3, columns 2 and 4)

| Indicator | World capacity that is BOF (%) | | | World capacity that is EAF (%) | | |
Scenario	BAU	Europe	All	BAU	Europe	All
World	71.6	70.5	56.5	28.4	29.5	43.5
North America	10.5	10.6	10.8	11.3	11.5	16.7
Europe	16.0	14.2	18.3	8.0	8.8	11.3
Asia	44.9	45.4	27.1	9.1	9.2	15.5

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia and Chile.

Table F.2.2: Simulation results with a dynamic panel data model and GMM (Table 5)

| Indicator | World capacity that is BOF (%) | | | World capacity that is EAF (%) | | |
Scenario	BAU	Europe	All	BAU	Europe	All
World	73.7	73.1	53.1	26.3	26.9	46.9
North America	10.8	11.2	14.5	10.4	10.7	18.9
Europe	18.2	15.8	28.9	7.7	7.8	14.2
Asia	44.1	45.5	8.8	8.1	8.4	13.7

Notes: North America includes Canada and the US; Europe includes Germany, Spain, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Spain, Portugal, Switzerland and Norway; Asia includes China, Japan, Russia, Turkey and India. Other countries included in the world average are Australia and Chile.

Table F.2.3: Simulation results using expected coal prices in levels (Table D.1.1) and the pre-sample mean estimator

| Indicator | World capacity that is BOF (%) | | | World capacity that is EAF (%) | | |
Scenario	BAU	Europe	All	BAU	Europe	All
World	71.6	68.6	53.5	28.4	31.4	46.5
North America	10.5	10.5	8.4	11.3	11.4	18.0
Europe	16.0	12.7	14.3	8.0	10.8	12.2
Asia	44.9	45.2	30.5	9.1	9.1	16.2

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia and Chile.

3. Heterogeneous effects across region and technology

Table F.3.1: Specifications with region-specific effects

Dependent Variable: share of units (x 1,000)	Pre-sample mean estimator BOF (1)	Pre-sample mean estimator EAF (2)	Dynamic panel data model BOF (3)	Dynamic panel data model EAF (4)
Log. expected coking coal price				
Europe	-0.00873	0.000938	-0.00923**	0.000388
	(0.00608)	(0.00126)	(0.00468)	(0.000910)
North America	-0.00885	0.000717	-0.0108**	-0.000113
	(0.00611)	(0.000761)	(0.00476)	(0.000657)
Asia	-0.00732	0.000838	-0.00863	0.000262
	(0.00716)	(0.000846)	(0.00562)	(0.000850)
Other	-0.00580	0.00324	-0.00621	0.000239
	(0.00696)	(0.00256)	(0.00480)	(0.00103)
Lag dependent variable	0.956***	0.963***	0.959***	0.970***
	(0.00515)	(0.00185)	(0.00708)	(0.00292)
Pre-sample mean (1970-1981)	-0.00337	-0.0183***		
	(0.00309)	(0.00274)		
Unit fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
p-value of Arrellano-Bond test for AR(2)			0.78	0.20
P-value of Hansen test			1.00	1.00
Degrees of freedom			307	317
Observations	14,506	26,888	18,286	33,794
Countries	19	18	21	21

Notes: The pre-sample mean estimator is estimated with 2SLS, and the 5th lag of the logarithm of expected coal price is used as instrument. The fixed effect dynamic panel data model is estimated with system GMM, and the instruments are built with the 3rd to 5th lags (of the logarithm of expected coal prices and the dependent variable). In all columns, cluster-robust standard errors in parentheses. Clusters are set at country level and *, ** and *** respectively denote significance at 10%, 5% and 1% levels.

Table F.3.2: Simulation results with region-specific effects and pre-sample mean estimator

Indicator Scenario	World capacity that is BOF (%) BAU	World capacity that is BOF (%) Europe	World capacity that is BOF (%) All	World capacity that is EAF (%) BAU	World capacity that is EAF (%) Europe	World capacity that is EAF (%) All
World	71.6	70.8	62.4	28.4	29.2	37.6
North America	10.5	10.6	10.3	11.3	11.4	14.4
Europe	16.0	14.6	17.1	8.0	8.6	10.1
Asia	44.9	45.3	34.6	9.1	9.2	13.0

Notes: North America includes Canada and the US; Europe includes Germany, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Portugal and Norway; Asia includes China, Japan, Turkey and India. Other countries included in the world average are Australia and Chile.

Table F.3.3: Simulation results with region-specific effects and a dynamic panel data model using GMM

Indicator	World capacity that is BOF (%)			World capacity that is EAF (%)		
Scenario	BAU	Europe	All	BAU	Europe	All
World	73.7	73.1	64.7	26.3	26.9	35.3
North America	10.8	11.0	11.4	10.4	10.5	13.3
Europe	18.2	16.9	21.6	7.7	8.1	10.4
Asia	44.1	44.7	31.0	8.1	8.2	11.5

Notes: North America includes Canada and the US; Europe includes Germany, Spain, Italy, France, Poland, Belgium, Czech Republic, Sweden, Finland, the Netherlands, Spain, Portugal, Switzerland and Norway; Asia includes China, Japan, Russia, Turkey and India. Other countries included in the world average are Australia (1982-1989) and Chile (2006-2014).

G. Unit root tests

We check for the stationarity of our dependent variable. We perform the Levin–Lin–Chu test, the Harris–Tzavalis test, the Im–Pesaran–Shin test and four Fisher-type tests to look for unit roots. The tests, reported below, reject the null hypothesis that the panels of our dependent variable have unit roots. This means that the assumption of stationarity in the econometric models appears to be valid.

We have also performed the Fisher-type tests for the independent variable (coal prices). We cannot perform the other tests since they require a strongly balanced dataset and we have missing data on that variable. The Fisher-type tests also go in the direction of stationarity.

Table G.1: Unit root tests

Variable	Share of installed capacity		Log. expected coking coal price	
Test result	Statistic	P-value	Statistic	P-value
Fisher-type unit-root tests				
- Inverse chi-squared	23,600	<0.0001	8701	<0.0001
- Inverse normal Z	-96.43	<0.0001	51.95	<0.0001
- Inverse logit L*	-110.14	<0.0001	50.12	<0.0001
- Modified inv. chi-squared Pm	154.07	<0.0001	55.09	<0.0001
Levin-Lin-Chu test	-14,000	<0.0001		
Harris-Tzavalis unit-root test	0.8868	<0.0001		
Im-Pesaran-Shin unit-root test	-33.84	<0.0001		

Notes: We subtracted the cross-sectional averages from the series before running the tests with the 1970-2014 data (since our models include year fixed effects). Fisher-type tests are based on augmented Dickey-Fuller tests, using one lag and with drift. The data was too unbalanced to perform the Levin-Lin-Chu, the Harris-Tzavalis and the Im-Pesaran-Shin tests on the log. expected coking coal price.