Relative Effectiveness of Energy Efficiency Programs versus Market Based Climate Policies in the Chemical Industry [1]

Gale A. Boyd

Corresponding author

Associate Research Professor

Social Science Research Institute, Duke University, Box 90989, Durham NC 27708

Jonathan M. Lee

Associate Professor

Department of Economics, East Carolina University, Greenville, NC 27858

April 28th, 2017

Abstract

This paper addresses the relative effectiveness of market vs program based climate policies. We compute the carbon price resulting in an equivalent reduction in energy from programs that eliminate the efficiency gap. A reduced-form stochastic frontier energy demand analysis of plant level electricity and fuel data, from energy-intensive chemical sectors, jointly estimates the distribution of energy efficiency and underlying price elasticities. The analysis obtains a decomposition of efficiency into persistent (PE) and time-varying (TVE) components. Total inefficiency is relatively small in most sectors and price elasticities are relatively high. If all plants performed at the 90^{th} percentile of their efficiency distribution, the reduction in energy is between 4% and 37%. A carbon price averaging around \$31.51/ton CO_2 would achieve reductions in energy use equivalent to all manufacturing plants making improvements to close the efficiency gap.

Keywords

Energy efficiency, price elasticities, manufacturing, stochastic frontier, plant-level data

[1] This paper was partially funded by the U.S. Energy Information Administration (EIA), Office of Energy Consumption and Efficiency Analysis while working with Leidos Inc. under contract with EIA (DE-EI0000564) and was prepared while the authors were Special Sworn Status researchers at the Triangle Research Data Center, a member of the Federal Statistical Research Data Center Network. The paper benefited from ongoing discussions and comments from Kelly Perl and Paul Otis on an earlier draft. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau or EIA. All results have been reviewed to ensure that no confidential information is disclosed.

Introduction

The Paris Accord has a goal of reducing CO_2 emissions to limit climate change to less than 2 degrees C. It has been broadly embraced by nearly every nation of the world. Even in the United States, where the Trump administration has stated its intent to pull out of the accord in 2020 and cancel the Clean Power Plan, climate policy is still being pursued by regions, states, cities, and via other existing programs at the Federal level. One might argue that the success of the Paris Accord is that it is not prescriptive regarding the types of policies nations must pursue, much the same way that a variety of policies are being pursued in the U.S., irrespective of the accord.

The policy options available to reduce CO_2 emissions are as varied as the sources of CO_2 themselves; de-carbonization energy use and reducing total energy demand are the overarching goals. The UNFCCC web site identifies 6 classes of policy instruments[2]; we consider three of these policy levers; market approaches, e.g. carbon taxes or tradeable permits; regulatory instruments, e.g. renewable portfolio standards or fuel economy / equipment standards; and voluntary programs, e.g. informational and behavior based interventions. Economists often point out the benefits of *market based approaches*, while others may point to market failures as a possible justification for the latter two instruments which we group together as *policies and programs (Jaffe and Stavins 1994)*. Market approaches rely on the price responsiveness of the demand sector to generate change in the level or mix of energy use. One important way that policies and programs work, particularly voluntary programs, is to reduce existing inefficiencies that may not be as responsive to price changes, per se. The relative effectiveness of market based approach (price) vs policies and programs (efficiency) will depend on the energy price elasticities of demand vis-à-vis the extent of existing levels of (in)efficiency. Sectors with high price elasticity and low levels of efficiency gap would be best tackled with market approaches; policies and programs would be more effective in sectors with the opposite, low elasticities and high levels of energy efficiency gap.

The extent and sources of the energy efficiency gap has been the subject of numerous studies and reviewed by a number of economists (Jaffe and Stavins 1994, Huntington 1995, Bloom, Genakos et al. 2010, Allcott and Greenstone 2012, Boyd and Zhang 2013, Boyd and Curtis 2014, Gillingham and Palmer 2014, Boyd 2016, Gerarden, Newell et al. 2017).

Jaffe and Stavins (1994) lay out a framework to compare the economic vs technical (engineering) perspectives in the early literature. Huntington (1995) draws a connection between a largely engineering approaches and the emerging data envelopment (DEA) and stochastic frontier analysis (SFA) literature measuring productivity, suggesting these economic tools could bridge the information gap. (Bloom, Genakos et al. 2010, Boyd and Curtis 2014) consider how the efficiency gap is related to management practices in industry, with mixed results. Bloom et al find a robust positive relationship in UK manufacturing data while Boyd and Curtis find no similar effect for US data. Allcott and Greenstone (2012) are critical of that

[2] http://unfccc.int/resource/climateaction2020/tep/policy-options/index.html accessed 7/5/2018.

largely engineering literature on the extent of energy efficiency and conclude that the area is ripe for rigorous study of how policies and programs might impact heterogeneous consumers. Gillingham and Palmer (2014) conclude the true sized of the gap is unclear, but also call for more research. (Gerarden, Newell et al. 2017) review the literature from the perspective of private vs. social optimality and market vs non-market failures. They find a wide range of evidence, from strong to weak, of both market and non-market failures, but note that this evidence is largely from the residential sector and less is known about the industrial and commercial sectors.

Most of these economic reviews recognize the important role of market based approaches, but few studies look at simultaneous price and efficiency effects and fewer still focus on the industrial sector. This paper addresses the basic empirical question with an industry case study of energy intensive chemical manufacturing in the U.S. Using stochastic frontier analysis (SFA) on the most detailed, plant-level data available, this paper econometrically estimates the two core elements needed for our comparison. The first is persistent and time varying energy efficiency gap, accounting for both industry-sector specific and plant level heterogeneity. The second is energy price elasticities, possibly accounting for energy price endogeneity. A two stage SFA is applied to estimate energy demand frontiers for electricity and fuel separately for 4 segments of the industry. This provides an estimate of the potential energy demand reduction in response to policies and programs designed to close the efficiency gap; the price elasticities can be used to assess a similar response to market based approaches. . While the cost-effectiveness of such programs are not considered, their potential impact is *assumed to be limited by the estimated, pre-existing efficiency levels*. We then compute the carbon price that would be needed to reduce demand by the same amount that is implied by the estimates of the energy efficiency gap. This price provides a metric to compare the two sources of carbon reductions.

The need for better information on industry energy efficiency was raised by Gerarden, Newell et al. (2017) and the importance of the industrial sector in terms of energy demand can't be overstated. In the U.S. 2017 Annual Energy outlook (U.S. Energy Information Administration 2017), industrial is the only sector that energy use is forecast to grow; residential, commercial, and transportation energy consumption are flat. This does not mean that industry doesn't respond to prices or experience technical change. While these types of improvements in efficiency occur in all sectors, improvements in the industrial sector are not forecast to keep pace with economic growth. It may even be the case that policies and programs in the non-industry sectors are easier to implement, e.g. CAFÉ standards in transport or appliance & lighting standards in residential and commercial sectors, and are responsible for the improvements in energy efficiency that have led to declining or stable demand. Regardless, it is likely that potential impact of energy policies and programs targeted at the industrial sector will be limited to reducing levels of current inefficiency, at least in the near term. The relative size of efficiency vs price response will be key to determining what policies are likely to be more impactful.

This paper provides estimates of energy efficiency and energy price response in the energy intensive chemical manufacturing sector and shares important features with the methodology presented in and applied to analyze metal based durables (Boyd and Lee 2016) in that it measures the distribution of energy efficiency of demand relative to local (plant level) energy prices. This paper proposes a slightly different approach from the MLE and LIML reviewed by (Amsler, Prokhorov et al. 2016), because we are concerned with *both price endogeneity and systematic plant level heterogeneity* in energy use which are not accounted for by Amsler et al. This paper uses a two stage variant of SFA that allows us to account for both plant level energy price endogeneity and plant specific heterogeneity in energy use. The two stage method developed by (Kumbhakar, Lien et al. 2014) handles heterogeneity can be readily modified to account for plant level price endogeneity controls. This modification is a contribution to the literature. The two stage approach allows the decomposition of efficiency into a plant specific (persistent) and time-varying components, which can be compared across new and continuing plants.   This is another contribution to the literature to explore the dynamic aspects of efficiency.

Jointly estimating the plant level price response and efficiency gap is important. For example, (Gerarden, Newell et al. 2017) point out that divergence of plant level price from average prices may overstate engineering estimates of the efficiency gap. Whether the energy efficiency measure from SFA or its non-parametric counterpart, Data Envelopment Analysis (DEA), is a purely technical efficiency measure or includes allocative efficiency depends on the inclusion of the relevant prices. (Filippini and Hunt 2015) discuss the difference in the treatment of technical and allocative efficiency in more detail. The energy sub-vector (directional) distance function as defined by (Boyd 2008) presents this approach as a measure of energy efficiency. Including prices makes the resulting measure of efficiency depend on prices but not explicitly measuring allocative efficiency by altering the direction in which energy efficiency is being measured. To illustrate this, Figure 1 shows the production isoquant for energy and all other quasi-fixed inputs and fixed output; the interior point A to the isoquant is inefficient. It shows three different directions that one could consider measuring efficiency, one is the purely technical efficiency measure in the absence of prices used by (Boyd 2008). The other two are based on different energy prices. The difference between the energy sub-vector (directional) distance function, AC, where the pure technical efficiency gap is measured as a reduction in energy use, holding other quasi-fixed inputs and production constant and one that includes both technical and allocative efficiency is embodied in the direction that efficiency is measured depending on prices. In this example, when one accounts for energy prices the efficiency gap is lower than when considering pure technical efficiency. In figure 1 at “high” relative energy prices the optimal energy use is E_1^* and the gap is E- E_1^* . In the presence of lower relative energy the level of the efficiency gap is smaller, E- E_2^*. We wish to estimate the plant level distribution of the efficiency gap, using local energy prices, and the price responsiveness represented by the own price elasticity.

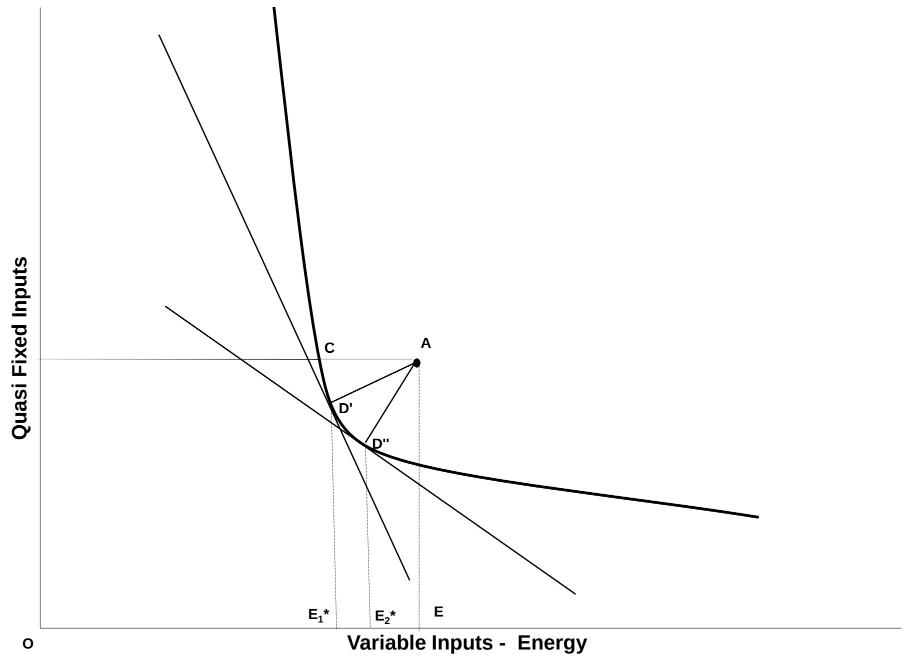

Figure 1 Comparison of Optimal Price Direction Distance Functions

Similarly, failure to account for sector and plant heterogeneity may also be sources of error in the efficiency gap estimation. We use detailed plant level data and can control the analysis at the 6 digit NAICS code level. Background on this confidential micro-data from the U.S. Census Bureau is presented in the next section. This NAICS level detail available in this dataset accounts for sector level heterogeneity in both the price and efficiency estimates. However, as described in more detail below, there is good reason to consider additional, within-sector, plant level heterogeneity. We model this using the two stage SFA.

We conduct a parallel analysis of two different sources of plant level data, as detailed below, since these data sources each have both advantages and disadvantages (see data section for overview). The use of these plant level data sources is another contribution to the literature. To our knowledge few, if any, studies of industrial energy demand and efficiency use plant level data. The closest we are aware of is (Bostian, Färe et al. 2016, Lundgren, Marklund et al. 2016, Zhang, Lundgren et al. 2016, Lundgren and Zhou 2017) all of which use either SFA or DEA on firm level data for Swedish manufacturing; (Lutz, Massier et al. 2017) using SFA on German firm level data. (Nguyen and Streitwieser 2008) (Bardazzi, Oropallo et al. 2015) estimate production functions from which price elasticities can be derived, but not energy efficiencies, using U.S. plant and Italian firm level data, respectively.

Once the estimates are obtained we conduct a simple policy exercise to compare the effectiveness of market vs efficiency programs for climate policy. We compute the carbon price that will reduce energy consumption by an amount equivalent to closing the energy efficiency gap at a level of performance equal to the 90th percentile of the estimated efficiency distribution. This is the final result and contribution of the paper.

This paper is organized as follows. The first sections provide an overview of energy intensive chemical manufacturing and describe the two plant level data sources at the core of the analysis. The next sections describe concerns over plant level energy price endogeneity and plant specific heterogeneity. A two – stage approach is presented as a solution. This section also introduces the notion of time-varying and persistent inefficiency. The parameter estimates for the elasticities and the distribution(s) for efficiency are presented for both data sources. Finally, the price (carbon or energy tax, etc.) that would be required to reduce energy demand, equivalent to the estimated levels of efficiency is computed. This provides a basis for comparison of possible relative effectiveness of market based or program based climate policy.

Background on the U.S. Chemical Industry

The Chemical Industry, as defined by the North American Industry Classification System (NAICS) code 325, is a diverse collection of sectors ranging from commodity chemicals (e.g. ammonia, chlor-alkalies, ethylene) to consumer products (e.g. paint, pharmaceuticals, cosmetics, etc.) The former are the up-stream process industries that encompass some of the most energy-intensive, chemical conversions of feedstock into intermediate chemicals, which are used primarily by other industries. The latter uses and produces a wide range of downstream chemicals to make, package and distribute final consumer goods. Of the over 5 Quads[3] of energy reported by the 2010 Manufacturing Energy Consumption Survey (MECS) that is used in NAICS 325, about 4.2 are used in the 13 energy intensive 6-digit NAICS listed below.

These energy intensive chemical sectors can be grouped into four chemical industry classifications that mimic the 4-digit NAICS hierarchical groups with some minor exceptions; Inorganic Chemicals, Organic Chemicals, Plastics and Resins, and Fertilizers (see table 1). This is the same industry sector grouping used by the National Energy Modeling System (NEMS) Industrial Demand Module (IDM) (Energy Information Administration 1994). This disaggregation of the chemical industry is important to more closely align the production activities into more homogenous groups for the purpose of energy analysis, as has been done in the NEMS model. However, even within these groups and associated 6-digit NAICS codes there is heterogeneity of energy use. For example, the primary chemical conversion in *NAICS 325311 - Nitrogenous Fertilizers* is from a feedstock, typically natural gas, to ammonia. Ammonia production is the most energy intensive step in the chain of nitrogen based fertilizer products. Ammonia is produced in a small number of plants and then used by other plants in the industry to manufacture other fertilizer products (Kermeli, Worrell et al. 2017). Ammonia may also be produced as an intermediate product in fully integrated production facilities. At the plant level there will be substantial differences in energy intensity even within this “narrowly defined” sector. There are similar examples in both organic chemicals and plastics & resins, where ethylene is the energy intensive primary chemical for a wide range of products. While the downstream products may also be energy intensive compared to other manufacturing, ethylene is much more so (Neelis, Worrell et al. 2008).

[3] This 5 quads includes feedstocks as well as energy for heat and power. All energy data are from the 2010 *MECS Table 1.2 First Use of Energy for All Purposes (Fuel and Nonfuel)*. These data are measured at end-use; i.e. electric generation losses are not included.

Table 1 NAICS Level Groups - Energy Intensive Chemical Sectors

NAICS	Industrial Sector
	Inorganic Chemicals
325120	Industrial Gases
325181	Alkalies and Chlorine
325182	Carbon Black
325188	Other Basic Inorganic Chemicals
	Organic Chemicals
325110	Petrochemicals
325192	Cyclic Crudes and Intermediates
325193	Ethyl Alcohol
325199	Other Basic Organic Chemicals
	Plastics and Resins
325211	Plastics Materials and Resins
325212	Synthetic Rubber
325222	Noncellulosic Organic Fibers
	Fertilizers
325311	Nitrogenous Fertilizers
325312	Phosphatic Fertilizers

Data

Data for the study are non-public plant-level Census Bureau data available in the Triangle Federal Statistical Research Data Center. These data are protected under Title 13 and 26 of the US Code and used with permission from the Bureau. Since these sectors are energy intensive, a parallel approach regarding the data sources is used. These data sources are the Manufacturing Energy Consumption Survey (MECS) and the quinquennial Census of Manufacturing (CM). MECS is a sample based survey conducted in 1985, 1988, 1991, 1994, 1998, 2002, 2006, and 2010[4]. The CM is part of the quinquennial Economic Census (EC); it includes all establishments operating during the analysis time period of 5 five-year time steps, 1992, 1997, 2002, 2007, and 2012. Both data span similar time periods but, for the most part, different years. The MECS and CM each have advantages and disadvantages which is why a parallel analysis approach was used.

Data needed for the analysis include energy use and prices along with production activities and other location specific variables. While the Manufacturing Energy Consumption Survey (MECS) provides the most detailed data on energy use, particularly cost and quantity of fuels by type, the MECS is a stratified sample and not a balanced panel so the presence (absence) of an

[4] 2014 was the most recent year, but not yet available to external researchers.

observation is not an indicator of entry (exit) in the industry. We need this information on entry/exit/continuing status for the relative efficiency of entering vs continuing plants. Using the Census of Manufacturing (CM), part of the quinquennial Economic Census (EC) solves this problem.

The availability of plant level electricity use and prices in the CM is one advantage of this data set. The CM provides plant level electricity consumption and costs, from which a plant level average price can be computed directly. (Davis, Grim et al. 2012) analyze the dispersion of those prices in detail. However, the CM only reports cost of fuels, not quantities, so Btu fuel consumption is imputed from fuel costs in the CM assuming state level average price of fossil fuels. (Boyd and Lee 2017) imputed fossil fuel use from the state average industrial price of natural gas. This was seen as a reasonable assumption for the metal based durables industries, because publicly available MECS data from 2010 for these 5 sectors suggests that 88% to 98% of the purchased fuel in this sector is natural gas. This is less true for energy intensive chemicals. MECS reports that in 2010 natural gas was only 77% of fossil fuels used for heat and power in these energy intensive sectors. This study imputes Btu consumption by taking the cost of fuels and dividing by a weighted average of the state level fossil fuel prices as published by the EIA's State Energy Data System (SEDS)[5], where the weights are computed from the published MECS data for each 6-digit NAICS above and applied to the closest year between the MECS and CM. Plants which generate part of their own electricity, not un-common in this industry, will likely purchase more fossil fuel and less electricity. To account for this the ratio of generated power to the total net consumption is computed.

Plant level shipment values, adjusted for inventory changes are used to measure production. Labor is measured in production worker hours. Capital stock computed in the micro data using perpetual inventory methods on investment data for both plant and equipment, separately (Foster, Grim et al. 2016). Non-energy material costs are computed by subtracting total material expenditures less cost for electricity and fuels. All data in dollar values are deflated using the NBER 6-digit NAICS price deflators (Bartelsman and Gray 1996). Sample statistic are shown below.

The MECS provides the most detailed data on energy use, particularly cost and quantity of fuels by type. The MECS is a sub-sample of the Census Bureau's Annual Survey of Manufactures (ASM) that targets mainly large plants[6]. MECS provides detail on a wide range of fossil fuel types, including the quantity of fossil fuels used as feedstocks. The plant level cost and quantity can be used to compute plant level average fossil fuel prices, as well as plant level electricity prices. While the MECS is a sub-sample its primary advantage is in the fossil fuel detail; a major component of energy use in this industry. MECS data on fossil fuel consumption is obtained by directly aggregating over all fossil energy types, excluding those used as feedstocks[7]. Costs are similarly aggregated and plant level average fuel prices are constructed and deflated to

---

[5] SEDS data is available online at the following: http://www.eia.gov/state/seds/ (last accessed November, 2016).
[6] In later years of the MECS the sample design is not strictly a sub-sample of the ASM, but we need data from the ASM on production and employment, so in those years we use the overlap between MECS and ASM.
[7] Data from the CM on fuel use states that this is for heat and power and should not include feedstocks, making these definitions comparable.

constant dollars using a GDP price deflator. MECS also indicates the amount of fossil fuels, mostly natural gas and gas liquids, as chemical feedstocks. Since plants using these feedstocks are likely to be more energy intensive, an indicator variable is created to reflect a plant is a feedstock using plant. All other economic variables in the MECS sample analysis are the same as those constructed for the CM.

Table 2 CM Sample Statistics (all values in logs, standard deviation below the mean)

Sector	Fuel	Electricity	Capital	Labor	Materials	Total value of shipments	Fuel Price	Electric Price
Inorganic	9.81	9.85	9.09	2.63	7.20	9.13	1.48	-2.93
	2.32	2.19	1.64	1.45	2.31	1.59	0.35	0.38
Organic	11.48	9.90	9.91	3.52	8.91	10.26	1.39	-2.9
	2.45	2.20	2.01	1.46	2.32	1.77	0.29	0.37
Plastic & Resins	10.11	9.56	9.33	3.48	8.34	9.86	1.48	-2.85
	2.41	2.14	2.01	1.43	2.56	1.78	0.29	0.36
Fertilizer	10.84	9.63	9.33	3.11	7.83	9.76	1.45	-2.9
	2.79	2.57	2.00	1.52	2.87	1.93	0.33	0.37
Total	10.43	9.75	9.41	3.20	8.10	9.73	1.45	-2.89
	2.51	2.21	1.92	1.50	2.53	1.78	0.32	0.37

Table 3 MECS Sample Statistics (all values in logs, standard deviation below the mean)

Sector	Fuel	Electricity	Capital	Labor	Materials	Total value of shipments	N. Gas Price	Electric Price	Other Fuels price
Inorganic	11.59	11.31	10.29	3.81	9.03	10.43	2.69	1.27	1.65
	2.58	2.09	1.49	1.37	1.99	1.44	0.41	0.46	0.93
Organic	12.86	11.32	11.06	4.39	10.66	11.43	2.73	1.29	1.65
	2.44	1.95	1.64	1.26	1.78	1.58	0.42	0.46	1.08
Plastic & Resins	11.81	11.42	10.82	4.57	10.72	11.4	2.72	1.29	1.76
	2.6	1.89	1.61	1.31	1.54	1.47	0.38	0.48	1.08
Fertilizer	12.79	11.36	10.63	4.32	10.36	11.44	2.69	1.13	1.84
	2.84	2.04	1.67	1.22	1.68	1.59	0.39	0.47	0.87
Total	12.11	11.35	10.7	4.25	10.13	11.1	2.71	1.27	1.7
	2.63	1.99	1.62	1.34	1.93	1.57	0.4	0.47	1.02

Methodology

This section briefly presents the ad-hoc demand model specification. This is done by adding energy prices to the energy factor requirement function described by (Boyd and Delgado 2012), which is equivalent to a directional input distance function. (Boyd and Lee 2016) motivate this by considering the energy prices as a modification of the direction of the distance function, but do not make that connection explicit. A review of stochastic frontier applications for energy use can be found in (Filippini and Hunt 2015). A modification of a two stage frontier estimation approach that handles heterogeneity is proposed as a solution that addresses both heterogeneity and price endogeniety in the first stage using instrumental variables (IV). Since the IV modification of the first stage of the frontier estimation is itself a two step approach we effectively have a three stage model. The last stage of the frontier approach also allows for the decomposition of efficiency into two components; one is plant specific and constant over time (persistent efficiency) and one that is time varying.

Stochastic Frontier Approach to Energy Demand

Following (Boyd and Lee 2016) we specify an SF ad hoc energy demand equation for the two primary energy types in each of the four sectors, with a few modifications, which are discussed below. We consider log linear models (KLEM Cobb-Douglas) of the general form,

$$lnE_{j,i,t} = f\left(lnY_{i,t}, lnK_{i,t}, lnEmp_{i,t}, lnNEM_{i,t}, lnP_{j,t,s}, DYear_t, DNAICS_k, GERATIO_{i,t}\right) + \varepsilon_{j,i,t} \quad (1)$$

Where

$lnE_{j,i,t}$	=	log of energy use
$lnY_{i,t}$	=	log of production or output
$lnEmp_{i,t}$	=	log employment or other measure of labor
$lnK_{i,t}$	=	log capital stock
$lnNEM_{i,t}$	=	log of non-energy material use
$lnP_{j,t,s}$	=	ln price of energy[8]
$DYear$	=	dummy for the year
$DNAICS_k$	=	dummy for the 6-digit NAICS code
$GERATIO_{i,t}$	=	ratio of self generated electricity to total purchased + generated - sold

j = energy type (electricity and fuel)

i = individual establishment (i.e. manufacturing plant)

s = state

t = year of the observations i.e. 1992, 1997, 2002, 2007, and 2012

k = 6-digit NAICS

[8] The subscript 's' refers to state level, but we use both state and plant level prices as detailed below.

The standard SF approach is to treat $\varepsilon_{j,i,t}$ as the sum two terms representing statistical noise, $v_{j,i,t}$, and inefficiency, $u_{j,i,t}$, respectively. We will return to specific approaches to the distributional assumptions of $\varepsilon_{j,i,t}$ below.

Total value of shipments (TVS), deflated and adjusted for inventory changes is used as the measure of productive output. Labor, measured by number of employees, controls for plant level utilization effects[9], since labor may be sticky in the short run. To better control for upstream and downstream plants within the sectors we include capital stock and non-energy materials. The most energy intensive chemical processes tend to be very capital intensive and have very simple material feed-stocks. Downstream plants may purchase chemicals produced by upstream plants and may involve simpler, less energy intensive production processes.[10] To account for this we consider non-energy material use. Non-energy material use, is the deflated costs of material purchases, less energy costs, which are included in the Census (material costs variables. The long run relationship between energy and plant scale is captured by the combined coefficient on production, capital, non-energy materials and labor. In a simple Cobb-Douglas specification the sum of the coefficients reflects the economies of scale with respect to energy. If the sum of the coefficients is less than unity then we can infer that larger plants will have lower frontier *energy intensity* than smaller plants. This means that the model will control for scale differences with respect to the energy efficiency measure.

Even within our 4 chemical sectors there can be a lot of heterogeneity of products and corresponding energy services, so 6 digit NAICS industry controls (industry fixed effects) are used in the empirical analyses. One could consider 10-digit product level dummies as well, since the CM has such detail. (Boyd 2016) reviews industry specific case studies of energy use that employ some of this finer product detail. However, doing so would require very specific prior information about which product level NAICS are more/less intensive, since there are a very large number of 10-digit product NAICS. We believe that the 6 digit controls are sufficient and are more detailed than other industrial energy studies have employed before. One exception is (Boyd and Curtis 2014) who also use plant level Census micro-data at the 6-digit level.

The price variables, $lnP_{j,t,s}$, reflects the impact of the prices of both electricity and natural gas on the frontier level of energy use. Incorporating prices into the factor requirement function allows us to measure price responsiveness of the sectors. If we view the model in a production function context then higher energy prices could act as an exogenous shifter of the frontier, i.e. induced technical change. The prices of both types of energy (j = electricity and fuel) may impact either energy type. Variation in energy prices can be used to capture price incentives and allocative efficiency. Electricity and fuel have different data issues in the CM, so the treatment of prices is different for the alternative energy sources. Specifically, Census data collects plant level cost and quantity for electricity but only costs for fossil fuels. The problem

[9] Using the 5-year Economic Census also conveniently avoids the years of the Great Recession by including 2007 and 2012, but not the intervening years.

[10] We estimate models with and without capital stock (see appendix for results excluding capital stock), because the capital stock variable is not available in our final CM year, 2012.

with using plant level electric prices[11] directly in the model is that the plant may have some bargaining power or simply more choice over rate plans, with larger electricity users realizing lower average prices, resulting in an endogenous variable.

We considered the possibility that Heating Degree Days (HDD) and Cooling Degree Days (CDD) could be used to control for ambient weather conditions on an annual basis using the zip-code location of the plant. Weather can impact building heating, ventilation and air-conditioning (HVAC) energy use, but also impact process energy via outside air to ovens and furnaces or chiller efficiencies, to the extent that the production requires these process. Preliminary analysis found these to be insignificant predictors of energy use with negligible impacts on plant efficiency. This is not surprising given the small role for HVAC in the chemical sector. As a result, these variable were excluded in the final model results.

Modeling electricity and fuel separately has advantages, since sector specific process needs will differ in terms of energy type. However, there may be opportunities to substitute electricity for fuel, combined heat and power (CHP) being the most obvious. Since Census data does include on-site generation we include a variable to control for this. We compute the ratio of self-generated power to the sum of self-generated power and purchased power minus sales to the grid and include it as a control variable. In the electric equation we would expect the generated electricity ratio coefficient to be negative (i.e. less purchased electricity), but in the fuel equation the coefficient would be positive account for the amount of extra fuel consumed in the CHP.

Directly estimating the model above faces some issues due to particular concerns in these sectors. These concerns are regarding endogeneity of energy prices and plant level heterogeneity that should be separated from efficiency. The next two sections describe these concerns, followed by our approach to account for them.

Plant Heterogeneity

Even within these 4 sectors of the chemical industry we anticipate plant level differences in processes and products that can require very different levels of energy. In organic chemicals the production of ethylene is much more energy intensive than subsequent downstream product. Ethylene is a component of many plastics, so if a plastics plant is fully integrated and produces its own ethylene then that plant would be much more energy intensive. Another example is ammonia production for fertilizers. This is a primary chemical input to other fertilizer chemical and is also produced as a final product. Ammonia production is a very energy intensive chemical to produce, but fertilizer plants may buy it instead of making it on site. There are other examples of producing sulphur related chemicals where the process is exothermic, i.e. since the reaction generates useable energy rather than requiring energy to sustain it.

One approach to account for plant heterogeneity would be use detailed material and product codes. This has been done by (Boyd and Delgado 2012, Boyd and Guo 2014, Boyd 2016) for

[11] These prices are not true marginal prices, but include demand charges, etc. They are total expenditures divided by total consumption.

some selected industries, but requires a large amount of knowledge regarding which specific material and product types are most relevant. Use of capital stock and material purchases might partially account for these plant level differences, since energy intensive plants are likely to have less expensive feed-stocks since they may make, rather than buy some intermediate product. Making an intermediate product is more likely to be both more energy intensive and more capital intensive as well. Even though we include capital stock and material purchases in the specification, additional methods to account for plant level heterogeneity are desirable.

The desire to distinguish between efficiency and heterogeneity requires an extension of the SFA frame work to a panel-data setting. The standard treatment for plant level heterogeneity in panel data is to include either a plant specific fixed or random effect. Equation (2) represents the non-stochastic frontier implementation of plant level heterogeneity by the inclusion of ω_i, for the i^th plant. ω_i, may be estimated by either a fixed or random effects estimator. In our application below we focus on results generated from a random effects estimator.

$$E_{i,t} = f\left(X_{i,t}\colon\theta\right) + \omega_i + \varepsilon_{i,t} \quad (2)$$

In the SF approach the typical error term is hypothesized to be made up of two parts,

$$\varepsilon_{i,t} = u_{i,t} + v_{i,t} \qquad (3)$$

Where $u_{i,t}$ is a one-sided efficiency error term and $v_{i,t}$ is noise. (Greene 2002) shows that this extension of the SF framework is econometrically tractable via maximum likelihood estimation (MLE). This approach has been labeled Greene's true fixed effect (TFE) and true random effect (TRE) estimators. In the TRE model, the estimates of ω_i are the basis for an estimate of persistent efficiency and $u_{i,t}$ is time varying efficiency. (Filippini and Hunt 2011, Filippini and Hunt 2012) employ this approach on panels of US states and OECD countries, respectively. However, these models can be difficult to obtain convergence in the MLE when the number of time periods is relatively small and the number of plants is relatively large. This was the same problem reported by (Boyd and Lee 2016) and is the case here as well. In this case the smallest plant sample size was 300 (Fertilizers), with most being over 2000. The number of time periods is five.

An alternative approach is to estimate these error components in a two stage process (Kumbhakar, Lien et al. 2014). The next section describes the two stage process. The advantage here is that the convergence problems are ameliorated and both heterogeneity and the price endogeneity can be treated in the first stage using a fixed effects for heterogeneity and an IV approach for price endogeneity.

Two stage model for persistent and time varying efficiency

The plant level efficiency estimates are obtained by a two stage approach. The first stage uses a plant level either fixed or random effects estimator with state level electricity prices as an instrument for plant level electricity prices. The general form for the estimate is

$$E_{i,t} = f\left(X_{i,t}\colon\theta\right) + \omega_i + \varepsilon_{i,t} \quad (4)$$

Where ω_i is the plant level fixed or random effect for the i[th] plant and $\varepsilon_{i,t}$ is Gaussian error. These two error components are not directly observable, but the residual of the regression, $E_{i,t} - f(X_{i,t}: \hat{\theta})$, can be decomposed into an estimate of the plant specific effect, $\widehat{\omega_i}$ that is constant over time for each plant and the time varying noise component, $\widehat{\varepsilon_{i,t}}$, based on the estimated parameters, $\hat{\theta}$.

$$\widehat{\omega_i} = E[\omega_i: E_{i,t} - f(X_{i,t}: \hat{\theta}), \hat{\theta}] \qquad (5a)$$

$$\widehat{\varepsilon_{i,t}} = E[\varepsilon_{i,t}: E_{i,t} - f(X_{i,t}: \hat{\theta}), \hat{\theta}] \qquad (5b)$$

The second stage is used to further extract efficiency estimates from the decomposed error terms using the stochastic frontier. Using the two plant level estimates from the first stage, a frontier analysis is conducted on each estimated error component

$$\widehat{\omega_i} = \alpha + u_i^{per} + v_i \quad (6a)$$

$$\widehat{\varepsilon_{i,t}} = \alpha + u_{i,t}^{tv} + v_{i,t} (6b)$$

Where the "usual" stochastic frontier model assumptions apply; u_i^{per} and $u_{i,t}^{tv}$ follow a one-sided exponential distribution and v_i and $v_{i,t}$ are noise. We are not interested in the estimate, $\hat{\alpha}$, per se, but in the estimates of $\widehat{u_{i,t}^{tv}}$ and $\widehat{u_i^{per}}$ based on the residuals, $\widehat{\omega_i} - \hat{\alpha}$ and $\widehat{\varepsilon_{i,t}} - \hat{\alpha}$, from each regression. The standard JMLS (Jondrow, Materov et al. 1982) frontier estimates from STATA of $\widehat{u_{i,t}^{tv}}$ and $\widehat{u_i^{per}}$ are obtained from these two 2[nd] stage regressions . The exponent of these JMLS estimates represent time-varying (*tv*) and persistent (*per*) efficiency.

$$tv_{i,t} = \exp(\widehat{u_{i,t}^{tv}}), \quad (7a)$$

$$per_i = \exp(\widehat{u_i^{per}}), \quad (7b) \text{ and}$$

$$\hat{}\, tot_{i,t} = \exp(\widehat{u_{i,t}^{tv}} + \widehat{u_i^{per}}) \ (7c).$$

Where $tot_{i,t}$ is the combined total efficiency estimate. We use a Hausman test to guide the choice of fixed or random effects in the first stage (equation 4).

Price endogeneity

Large energy users, either by virtue of sheer size or by virtue of having energy intensive production processes have good reasons to get the lowest possible energy prices. This means that lower plant level prices would be correlated with higher energy demand for reasons other than pure price responsiveness, i.e. estimated price elasticities would be biased upwards in absolute terms due to simultaneity bias. It is possible that concerns over endogeneity are adequately captured by the fixed (or random) effects approach used to capture heterogeneity, as described above. However, the two stage approach gives us the ability to additionally explore endogeneity concerns. The two data sets constrain our explorations; we were limited to electricity use in the CM analysis, which includes data on plant level electric use and price but not fuel prices. The MECS data analysis allow for analysis of both plant level electricity and fuel price endogeneity to be considered as well.

Concerns regarding and methods to control for endogeneity in the maximum likelihood SF context is reviewed by (Amsler, Prokhorov et al. 2016). One approach is the control function. In this method, plant level energy prices are regressed against the instrument, in this case state level price, and all the independent variables of the SF model. The residuals of this first stage regression are included in the SF estimation. A significant coefficient on the residuals indicates the prices are endogenous. We propose a different approach using IV. We employ IV estimators to instrument plant level prices with average state level industrial energy prices in the first stage of the two stage SFA. When models are linear in parameters the IV and CF approaches are expected to provide similar results and the two stage frontier method we use here makes the IV easy to implement in the first stage by modifying equation (4). State level prices are published by the Energy Information Administration (EIA) State Energy Data System (SEDS). The advantage of these data is that they are collected consistently by EIA over the time period from surveys of the utility service companies rather than from the plants. The average price charged to industrial consumers will account for both the cost structure of the utility, regional/temporal differences in utility regulation/restructuring, and the different rates structures available to all industrial consumers. We conduct a Hausman test to determine the validity of our IV in the first stage (equation 4).

Empirical Results

The model and estimation approach described above is applied to a panel dataset for the CM and MECS in separate analysis. In principle, these data sets could be pooled, but we do separate analyses for two reasons. The first is that the MECS is a stratified sample (over-sampling energy intensive plants) and the CM is a Census, i.e. includes all plants. We wish to explore how these two data collections might impact the results. The second is that MECS has much more detail on energy, including physical measure of fossil fuels and the corresponding detail needed to compute fuel specific, plant-level prices. The detailed nature of the MECS also might result in different persons within a firm/plant to be tasked to fill out the survey form, compared to the CM. In some sense the MECS might include better or more accurate data on energy use, particularly with respect to fuels.

The next section presents the results from the CM analysis for each of the four sectors and two energy types. The impact of the IV approach is discussed. The efficiency estimates are discussed in some detail, including the decomposition of efficiency into persistent and time varying and the comparison of efficiency for existing and new plants, i.e. whether new plants that enter the industry are more efficient than their counterparts. Finally, we explore the aggregate implications for the estimated distribution of total efficiency in these sectors.

The subsequent sections highlight some differences that arise from using the MECS sample. These include the ability to instrument for fossil fuel prices using the same approach as employed in the CM dataset.

Two Stage SFA Parameter Estimates – CM data

In the first stage of the two-stage process the ad hoc energy demand model described in equations (1) and (2) using both fixed and random effects for ω_i were estimated. A Hausman test rejects the null hypothesis of RE in all cases over FE (95% confidence level), i.e. found that the FE was preferred over the plant level random effects. Using the FE model for heterogeneity, we instrument for endogeneity of plant level electricity prices using state level prices inform the EIA SEDS data. A Hausman test for the IV does not reject the null hypothesis. To the extent that plant level prices show systematic variation as described above, it appears that the FE treatment for plant level heterogeneity is adequate to deal with possible price endogeniety as well. Two sets of analysis were done; one includes capital stock and the other does not. Capital stock data is not available in the CM for 2012. Estimates of the price elasticities and efficiency measure are very similar[12].

Tables 4 show the estimates for each industry and energy type. Own price elasticities range from -0.7 to -1.2. Non-energy materials is negative and significant in inorganics for electricity use. Inorganic chemicals is a very diverse collection of products and processes, some of which are quite electric intensive. One is industrial gases. It may be that some plants in this sector primarily mix or bottle gases made elsewhere for delivery. If that is the case, then those plants might have high non-energy material shares. Examining the micro data suggests that this was often the case; this sector had the highest level of variation in non-energy material shares. Even with the 6-digit NAICS control we believe that the negative and significant coefficient reflects this underlying phenomenon. Non-energy materials is negative and significant for fuel use in fertilizers, but the reason is less clear.

The self-generation ratio is generally significant and has the expected sign, positive for fuel and negative for electricity, with exception of fuel use in fertilizers. The mean time varying and persistent efficiencies range from 0.7 to 0.98 and most have small standard deviations. Since total efficiency is the product of the two components the mean for total efficiency is smaller, ranging from 0.58 to 0.87. We will take a closer look at the efficiency distributions in the next section.

[12] Results for the models without capital stock, but including 2012, are available on request.

Table 4 Two Stage SFA Estimates, by Industry and type of Energy, CM data

	Inorganic		Organic		Resins & Plastics		Fertilizer	
	Electricity	Fuels	Electricity	Fuels	Electricity	Fuels	Electricity	Fuels
$lnEmp$ Log Employment	0.105	0.270***	0.184***	0.109	0.231***	0.113	0.295	0.291
lnK Log Capital	0.274***	0.195***	0.145***	0.344***	0.213***	0.240***	0.153	-0.0801
$lnNEM$ Log non-energy Materials	-0.195***	-0.0131	0.0322	0.0199	-0.032	0.0321	0.00257	0.489**
lnY Log Total Value of Shipments	0.710***	0.467***	0.466***	0.297***	0.508***	0.433***	0.464**	0.162
$GERATIO$ Self Generation Ratio	-0.218	2.311***	-2.318***	0.679***	-1.424	1.727***	-1.370***	-0.713**
lnP_{NG} Log Natural Gas Price	-0.136	-1.179***	0.246**	-0.688***	-0.0561	-1.004***	-0.32	-1.033*
lnP_{Elec} Log Electricity Price	-0.965***	-0.0972	-0.704***	-0.178	-0.890***	-0.108	-1.206***	-0.399
Constant	-0.36	4.373***	0.345	4.737***	0.388	3.821***	-0.0971	4.408
Observations	2400	2400	1900	1900	2300	2300	300	300
Number of Firms	1400	1400	1100	1100	1300	1300	200	200
Time-varying Efficiency	0.842	0.789	0.868	0.817	0.893	0.818	0.888	0.785
Std Dev	0.078	0.0783	0.0636	0.102	0.0461	0.0352	0.0665	0.373
Persistent Efficiency	0.689	0.938	0.752	0.953	0.976	0.963	0.966	0.773
Std Dev	0.0882	0.147	0.0752	0.114	0.0697	0.0871	0.0492	0.0807
Overall Efficiency	0.58	0.74	0.653	0.778	0.872	0.787	0.857	0.609
Std Dev	0.0915	0.137	0.0682	0.0811	0.0592	0.0853	0.0522	0.3

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Including capital stock allows us to look at the short and long run effect of scale on energy use. We define the short run elasticity of scale relative to the demand function

$$lnE_{j,i,t} = f\left(ln\lambda_s Y_{i,t}, lnK_{i,t}, ln\lambda_s Emp_{i,t}, ln\lambda_s NEM_{i,t}, \dots\right) + \varepsilon_{j,i,t} \qquad (8a)$$

And long run relative to

$$lnE_{j,i,t} = f\left(ln\lambda_l Y_{i,t}, ln\lambda_l K_{i,t}, ln\lambda_l Emp_{i,t}, ln\lambda_l NEM_{i,t}, \dots\right) + \varepsilon_{j,i,t} \qquad (8b)$$

i.e. in the short run only output, employment and non-energy materials are variable, in the long run capital is also variable. The short and long run elasticies are

$$\frac{\partial lnE_{j,i,t}}{\partial \lambda_s} = \beta_Y + \beta_{Emp} + \beta_{NEM} \qquad (9a)$$

$$\frac{\partial lnE_{j,i,t}}{\partial \lambda_l} = \beta_K + \beta_Y + \beta_{Emp} + \beta_{NEM} \qquad (9b)$$

In other words, the long run elasticity of energy with respect to scale as the sum of the estimated coefficients, β_* for labor, non-energy materials, total value of shipments, and capital stock. This would measure the percentage impact on energy use of a larger plant for a percent change in both variable and fixed inputs and outputs. In the short run, capital is fixed. The short run elasticity is the sum of the coefficients for the variable inputs and output. It may be best thought of as the elasticity of plant utilization, given fixed capital. Table 5 shows that, in the long run, the elasticity of scale is close to, or slightly less than unity. The smaller short run elasticities reflect that as variable inputs and production fall, relative to the fixed capital stock, energy use falls less than proportionally[13]. This is consistent with observations that the energy output ratio tends to rise as plants produce at less than full capacity over the business cycle.

Table 5 Elasticities of Scale with respect to Energy Use

	Inorganic	Organic	Resins and Plastics	Fertilizer
	Electric			
Long Run	0.89	0.92	0.92	0.91
Short Run	0.62	0.72	0.71	0.76
	Fuel			
Long Run	0.92	0.77	0.92	0.86
Short Run	0.72	0.43	0.71	0.94

Efficiency Results - Census of Manufacturing

This section explores the efficiency estimates from the CM analysis in more detail. The mean and standard deviations for the plant level efficiencies show that the level of efficiency is, for the most part, fairly high and tightly distributed. For electricity, resins & plastics have the highest efficiencies, followed by fertilizers, organics, and inorganic chemicals; the range is from 0.58 to 0.87. For fossil fuels resins & plastics are the most fuel efficient, followed by organics, inorganics and fertilizers; the range is from 0.61 to 0.79.

One consideration is whether new plants that enter the industry might be more efficient than their existing counterparts. For example, a new plant can have more advanced technology, but may initially exhibit poor operations management. Over time, with learning, a new plant may become more efficient. We compare the mean time-varying (TV) and persistent (PER) efficiency of new vs existing plants in table 6. A pattern emerges. In 6 out of 8 sector-energy combination the time varying efficiency of new plants is worse (lower) and statistically significant via a t-test for difference in group means. The exception is fuel use in fertilizer. The pattern is opposite for persistent efficiency, except for electricity use in organics and Resins & Plastics. We interpret this as that while new plants may have slight advantages in technology, when they first enter the industry those advantages are not fully realized, i.e. start out with lower time-varying efficiency. Over time this difference in time-varying efficiency goes away, i.e. these plants learn by doing as they become existing plants five years later (the next CM year of the data). This analysis does not explore these dynamics in detail, so this is a hypothesis to examine in future research.

[13] The one exception to this pattern is for fuel use in fertilizer, where the capital stock estimate is negative and insignificant. In this case we would anticipate that both the short and long run estimates are the same.

Table 6 Comparison of Efficiency of New vs Existing Plants by sector and energy type

		Inorganic			Organic			Resins & Plastic			Fertilizers		
		New	Existing	t-test	New	Existing	t-test	New	Existing	t-test	New	Existing	t-test
Electric	TV	0.82	0.83	-2.17	0.86	0.86	-2.45	0.87	0.88	-3.92	0.88	0.89	-0.82
	PE	0.96	0.96	1.68	0.82	0.83	-6.74	0.82	0.83	-9.02	0.97	0.96	1.65
Fuel	TV	0.79	0.79	-2.24	0.79	0.81	-4.11	0.80	0.81	-3.70	0.76	0.77	-0.14
	PE	0.93	0.92	0.87	0.96	0.95	2.35	0.96	0.96	1.88	0.84	0.75	2.46

The kernel densities for overall efficiency, shown in figures 2 and 3, reveal more about total efficiencies. First, the distributions for fossil efficiency are more tightly clustered than for electricity, but more importantly there are virtually no plants that might be called "highly inefficient"; the left tail is very thin. In addition, there are few plants that are considered 100% efficient; this puts the mean efficiency estimates into a different light. Mean efficiency ranging from 0.58 to 0.87 might suggest opportunities for average energy reduction by anywhere from 13 to 42%, relative to an absolute efficiency of 100%. Strictly speaking that is true, but the distributions suggest that 100% efficiency isn't common. Another way to view the overall level of efficiency in each industry would be to compute the reduction in aggregate energy use if all plants were "efficient", i.e. achieve some empirically relevant level of performance other than 100%. Since there are empirically few plants that are 100% efficient, we define an "efficient plant" as one that performs at the 90th percentile of the corresponding efficiency distribution. We take the plant level energy use and reduce it by the amount needed to put it at the 90th percentile. If the plant is already at the 90th or greater, then the plant is already "efficient." The ratio of the sum of the "efficient" energy consumption to the sum of the actual energy use reflects the potential level of energy use if all plants were efficient. This might reflect policies and programs that achieve improvements at a high level of comparative performance, the 90th percentile, but not perfect performance, 100% efficiency. This becomes the basis for our comparison of market based vs program based climate policies.

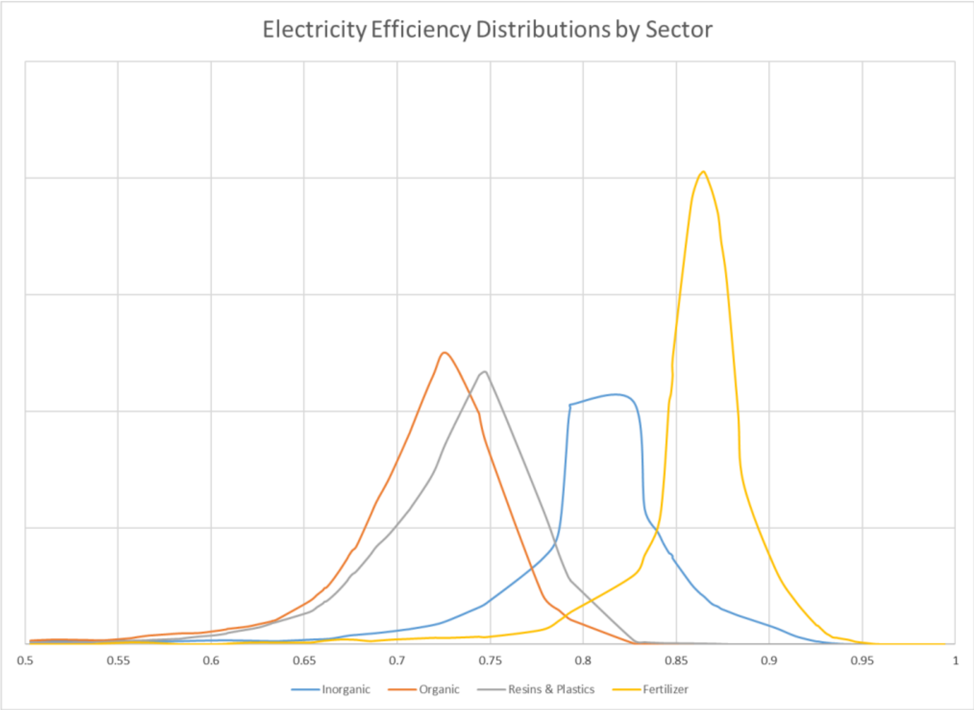

Figure 2 Kernel Density for Plant level Electricity Efficiency by sector

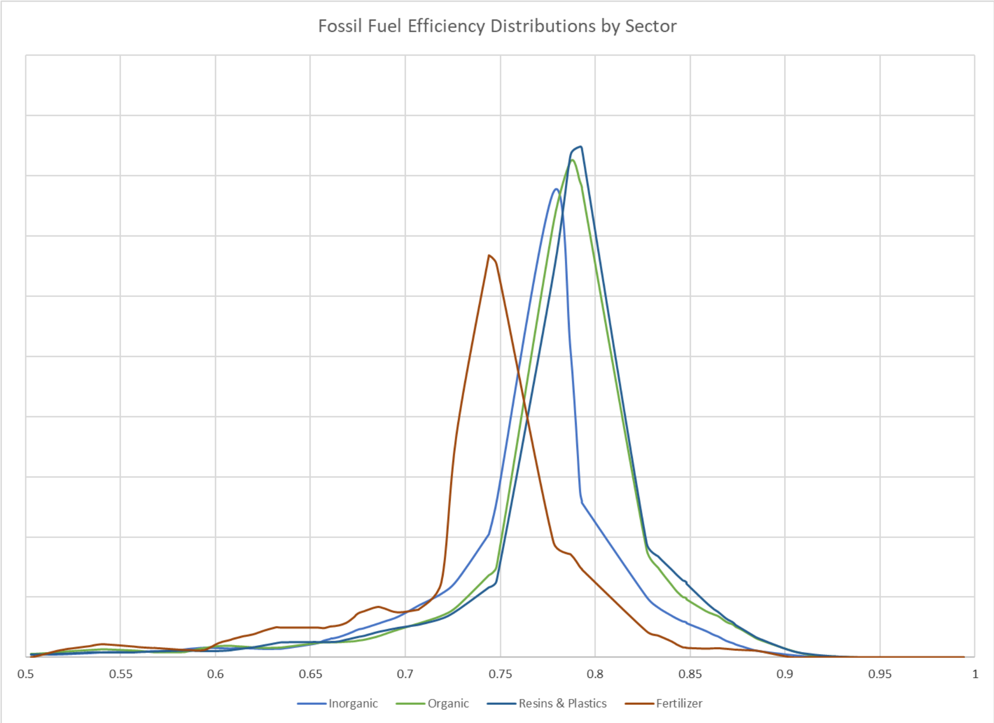

Figure 3 Kernel Density for Plant level Fossil Fuel Efficiency by sector

Panel A of Table 7 shows the potential percent reduction in energy from eliminating inefficiency as measured by one minus the above ratio. While the analysis suggest that substantial gains are possible for electricity use in inorganic and organic chemicals, opportunities are low elsewhere. The average is about -8% elsewhere. This may seem like a small percentage, but since the base level of energy use in this sector is large this is a substantial amount of energy reduction. This result is also consistent with what (Boyd 2016) reports in a meta-analysis of 2 dozen industry case studies; energy intensive sectors have much tighter distributions of estimated efficiency than non-energy intensive sector. It is likely that in competitive industries that produce energy intensive commodities markets will not tolerate as much energy inefficiency.

Table 7 Potential Reduction in Energy Use if All Plants Were Efficient (90th percentile) by sector and energy type

Panel A: Percentage Change in Energy Use at Frontier				
	Inorganic	Organic	Resins & Plastic	Fertilizers
Electricity	-37%	-28%	-8%	-4%
Fuels	-9%	-7%	-9%	-8%
Panel B: Carbon Price (\$/Ton CO_2) Achieving Equivalent Reduction				
	Inorganic	Organic	Resins & Plastic	Fertilizers
Electricity	58.39	153.17	10.57	2.93
Fuels	6.46	7.67	7.50	5.38

Panel B of Table 7 considers an alternative policy instrument designed to achieve an equivalent reduction in energy use as the percentages reported in Panel A. Specifically, the equivalent tax on CO_2 emissions is calculated from the parameter estimates in Tables 1-4 evaluated at the US average prices for electricity (\$19.6/MMBtu) and natural gas (\$4.91/MMBtu) in the industrial sector in the most recent 2012 wave of our panel. The formula used to derive the required carbon tax (CT) for electricity reductions is given by the following:

$$\%\Delta = \frac{(\$19.6+CT*0.1667)^{\beta_{ee}}(\$4.91+CT*0.0597)^{\beta_{ef}}}{(\$19.6)^{\beta_{ee}}(\$4.91)^{\beta_{ef}}} - 1. \quad (10)$$

The left-hand side of equation (10) is set equal to the energy reductions calculated in Panel A of Table 6, β_{ee} and β_{ef} are the own price elasticity for electricity and the cross price elasticity for fuel (from the IV models in Tables 1-4). The fixed parameters 0.1667 and 0.0597 are the CO_2 emissions factors (tons CO_2/MMBtu) for electricity purchases and natural gas, respectively.[14] The formula for fuel simply uses the corresponding elasticities from the fuel estimates. The weighted-average carbon tax is \$ 36.61/ton CO_2 would achieve equivalent reductions in energy

[14] Average US industrial electricity and natural gas prices are from the US EIA's State Energy Data System (SEDS). CO_2 emissions factors are calculated from the EPA's Emissions & Generation Resource Integrated Database (eGRID) for 2012. SEDS data is available for download at the following: https://www.eia.gov/state/seds/ (last accessed January, 2018). eGRID data is available online: https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid (last accessed January, 2018).

use in the chemical sector as compared to requiring all manufacturing plants make efficiency improvements to reach the frontier. To put this number in perspective, the US EPA and the General Accounting Office currently estimate a median value of the social cost of carbon equal to \$31.84 per ton CO_2 or roughly four times the amount of carbon tax required to achieve equivalent energy reductions in comparison to efficiency investment requirements (United States Government Interagency Working Group on Social Cost of Carbon 2013, US EPA 2015).

However, there is substantial variation in the sector specific carbon taxes. In particular, much higher estimates of electricity inefficiency results in high carbon tax equivalents. The simple average of the carbon tax equivalent for the remaining 6 simulated values is only \$6.75. If we compare these carbon prices to the European Union carbon market we see that for the last six years a ton of CO_2 was trading well under €10 (\$11.30), but has risen recently above €15 (\$16.95).[15] Some projections of the EU carbon market are for it to go higher in the future[16]. These results suggest that a tax on carbon emissions computed above is well within range of market outcomes in the EU and lower than U.S. estimates of the social cost of carbon, so a market instrument is likely to be preferable to alternative to technology mandates for reducing energy use in the sense that similar outcome would be achievable at a relatively low price, at least in this particular energy intensive sector.

One concern about this simple simulation is that policies and programs, particularly efficiency standards, may be subject to rebound. Rebound is the theory that when an energy service is made more efficient it is effectively becoming cheaper. If a new technology is 10% less energy intensive the user of the energy service might demand more, resulting in a smaller amount of energy being "saved." If the price elasticity of demand for the energy service is greater than unity, the rebound effect can result in more energy being consumed. In a study of Swedish energy intensive industry that also uses an SFA approach (Amjadi, Lundgren et al. 2018) find a substantial level of rebound, such that a 20% efficiency reduction target policy may only result in 5% to 8% savings, for chemical sector specifically. If the U.S. chemical has a similar level of rebound then in the context of our simply simulation, the effort (in the form of an efficiency target) might need to be much higher than percentage changes in table 7. Interestingly, the 20% policy scenario by Amjadi, Lundgren et al results in level of percentage savings in a similar range to ours. Our computation of a carbon price is based on the *outcome* of some unspecified policy, not the target. While the possibility of rebound may imply that a policy might need to target a higher level of efficiency, similar to the scenario in Amjadi, Lundgren et al, to achieve the reductions we use in the simulation, the corresponding carbon price based on the final result of some policy or program will be unchanged.

---

[15] Conversion using 2017 average exchange rates.

[16] **Reuters Environment** APRIL 11, 2018 / 7:22 AM / *Analysts raise EU carbon price forecasts on emissions rise, UK Brevity clarity* https://www.reuters.com/article/us-eu-carbon-survey/analysts-raise-eu-carbon-price-forecasts-on-emissions-rise-uk-brevity-clarity-idUSKBN1HI1LR accessed 7-19-2018

MECS Two Stage Parameter Estimates

We conduct a parallel analysis to that employing the CM data, but using the MECS sample. The major difference is that the MECS data allows us to use plant level fuel and electric prices, with fuel prices begin further broken down into natural gas and all other fuels. The use of plant level fuel prices raises the same price endogeneity issues as the plant level electric prices does in the CM data. The same two-stage estimation strategy is used, but we instrument the plant level energy prices with the corresponding state level prices for electric, natural gas, and other fuels. The Hausman test for FE vs RE are the same as for the CM results; FE models are preferred in all cases and at high levels of significance. The same result applied for the Hausman test for IVFE vs FE; FE is preferred over the IV. Similar to the CM analysis, FE seems adequate to handle any concerns over price endogeneity.

Results are shown in table 8. Despite the notion that the MECS data may be more accurate, given the level of detail, the price elasticity results are not as satisfying. Many of the coefficients are not significant; others are significant but the wrong sign.

Table 8 Two Stage SFA Estimates, by Industry and type of Energy, MECS data

	Inorganic		Organic		Resins & Plastics		Fertilizer	
	Electricity	Fuels	Electricity	Fuels	Electricity	Fuels	Electricity	Fuels
$lnEmp$ Log Employment	0.367***	0.259	0.204***	0.262***	0.287***	0.315**	0.132*	-0.0095
lnK Log Capital	0.254***	0.315***	0.386***	0.543***	0.348***	0.699***	0.579***	0.17
$lnNEM$ Log non-energy Materials	-0.028	0.0862	0.00594	-0.0011	0.0326	0.103	0.0324	0.0539
lnY Log Total Value of Shipments	0.0567	0.147	0.115**	0.116	0.176	0.181	-0.00163	0.0519
$GERATIO$ Self Generation Ratio	D	D	D	D	D	D	D	D
lnP_O Log Other Price	-0.00966	-0.125**	0.00791	0.0558	-0.0171	0.00601	-0.00484	-0.0221
lnP_{NG} Log Natural Gas Price	0.0158	0.509***	-0.0526	-0.186	-0.0219	-0.0825	0.309**	0.061
lnP_{Elec} Log Electricity Price	0.495***	0.490*	-0.204**	0.261	0.569***	-0.523**	-0.364	-0.327
Constant	7.356***	4.352***	6.433***	5.430***	5.862***	3.44	7.862***	11.51***
Observations	1300	1300	1100	1100	1300	1300	300	300
Number of firms	700	700	600	600	500	500	100	100
Time-varying Efficiency	0.902	0.768	0.932	0.815	0.917	0.763	0.915	0.853
Persistent Efficiency	0.956	0.903	0.961	0.938	0.973	0.943	0.936	0.777
Overall Efficiency	0.862	0.694	0.895	0.766	0.892	0.72	0.856	0.663

*** p<0.01, ** p<0.05, * p<0.1 , D = withheld for disclosure purposes

Summary

This paper presents estimates of the distribution of energy efficiency and price elasticities in the four major energy using sectors of the upstream, energy-intensive portions of the Chemical Industry. We analyze data from the CM and MECS separately, since these data sources have their own strengths and weaknesses. If we compare the mean efficiency estimates between the two data sets (table 9) the mean fuel efficiency are fairly similar. Electricity efficiency in inorganics and organics differ the most. There is no evidence of bias, in the sense that one data source uniformly has higher or lower mean efficiency. Most of the mean efficiency differences are not particularly large, except electricity in two sectors.

The CM analysis, since it is not a sample, is the preferred source for the aggregate analysis of the potential savings from efficiency, since all plants are included in the data. That analysis shows that the range of efficiency difference is for the most part, quite narrow and the total savings associated with moving all inefficient plants to the 90th percentile of the frontier distribution is small in percentage terms, ranging from a low of 4% to a high of 9%, depending on the sector and energy type, with the exception of inorganic and organic electricity use. When we look at the details, we see that persistent inefficiency is the main source of this result. Time varying efficiency is similar across the sectors and energy types. The relatively small percentage difference in efficiency for 6 out of the 8 energy sector combinations is consistent with other studies that find energy intensive sectors e.g. steel, cement, paper, etc. (Boyd and Zhang 2013, Boyd and Guo 2014, Boyd, Doolin et al. 2017) have a much narrower range of efficiency than less energy intensive ones, e.g. metal based durables, auto assembly, etc. (Boyd 2014, Boyd and Lee 2016). We find that new plants have slightly higher persistent efficiency than existing plants, but enter the industry with lower time-varying efficiency. We interpret this as a new plant learning phenomenon, but this analysis doesn't model this explicitly. The results for fertilizers might bear further examination since this sector was the most sensitive to the data source (CM vs MECS) and model specification.

Table 9 Comparison of Mean Efficiency Estimates

	Inorganic		Organic		Resin & Plastics		Fertilizer	
	Electric	Fuel	Electric	Fuel	Electric	Fuel	Electric	Fuel
MECS	0.862	0.694	0.895	0.766	0.892	0.720	0.856	0.663
CM	0.580	0.740	0.653	0.778	0.872	0.787	0.857	0.609

When comparing price elasticities (table 10), estimates from MECS are all lower and many are not significant. MECS is a sample with much more detail about different types of energy use; the CM is the entire universe of plants and have general more observations in the modeling analysis. Using the lower MECS price elasticities would impact the analysis presented in the simulation, but the lack of precision in the MECS estimates led us to prefer the CM for the analysis.

Table 10 Comparison of Own Price Elasticity Estimates

	Inorganic		Organic		Resin & Plastics		Fertilizer	
	Electric	Fuel	Electric	Fuel	Electric	Fuel	Electric	Fuel
MECS	-0.495***	-0.509***	-0.204**	-0.186	-0.569***	-0.0825	-0.364	0.061
CM	-0.965***	-1.179***	-0.704***	-0.688***	-0.890***	-1.004***	-1.206***	-1.033*

The estimate of the efficiency gap, defined not by 100% efficiency but by the 90th percentile of the estimated plant efficiency distribution, is relatively small. Since many of the CM price elasticities are near unity, this results in a rather modest carbon price equivalent to closing the efficiency gap of \$6.75/ton CO_2 for all but electricity use in two sectors. When those are included the average is \$31.51/ton CO_2 . The MECS analysis does generate lower elasticities, but similar efficiency estimates and the resulting equivalent carbon prices only rise to \$14.01/Ton CO_2 (weighted average). All are below the social cost of carbon of \$31.84/Ton CO_2. While the effectiveness of programs targeting the efficiency gap depend on many things, this analysis suggests that only very modest carbon price would be needed to get similar impacts, with the possible exception of electricity use in two sectors. Since the higher levels of inefficiency are driven by the estimate of persistent efficiency, expecting to achieve those levels in practice might be questionable. If programs are easy to implement or have positive synergies then both policies might still be pursued, but the modest nature of the tax to generate energy reductions suggest that, at least for this sector, market approaches may be more effective.

Conclusions

This paper uses a detailed plant level micro-data set to analyze four energy-intensive sectors of the chemical industry. Applying a frontier estimation method to obtain estimates of both price elasticities and levels of efficiency we conduct a very stylized simulation to compute the equivalent carbon tax that would "eliminate" the estimated inefficiency. The result is a modest carbon tax that is similar current market prices and social cost of carbon estimates. The analysis cannot conclude if efficiency programs, regulations, or other non-market approaches would be more, or less cost effective than a tax. It also does not measure if there are possible rebound effects of such programs or synergies with a carbon tax.

One important observation is, in principle, 100% efficiency should be possible but the empirically estimated efficiency distribution may show that high levels of efficiency aren't common. Simply looking at the mean efficiency may not tell the whole story, empirically. A more meaningful level of efficiency might be based on some percentile level of the estimated probability distribution. For our purposes, 90th percentile is used but this is an arbitrary choice since the analysis can't assess the effectiveness of program to reach any particular level of efficiency.

One extension of this analysis would be to extend this analysis to a wider range of industries. Another would be to follow related work using SFA to estimate such rebound (Adetutu, Glass et al. 2016). Examining how participation in various state, local or federal efficiency programs impacts the position of plants/firms within the distribution would be very interesting. (Dalzell, Boyd et al. 2017) developed a data set that links plants that received energy audits and found that they do move up in the efficiency distribution, but that more work was warranted. (Boyd and Curtis 2014) find that “good management” of firms has a mixed impact on plants within their industry specific efficiency distribution. Some practices were beneficial and others detrimental to improved energy performance. That work was based on correlations and not “causal” analysis, like a difference in difference. The fact that the CM micro-data has a relatively long time span of plants and firm ownership provides the basis for expanding this type of analysis; developing robust data on management and program participation is the bigger challenge.

REFERENCES

Adetutu, M. O., A. J. Glass and T. G. Weyman-Jones (2016). "Economy-wide Estimates of Rebound Effects: Evidence from Panel Data." The Energy Journal **37**(3): 251-269.
Allcott, H. and M. Greenstone (2012). "Is There an Energy Efficiency Gap?" The Journal of Economic Perspectives **26**(1): 3-28.
Amjadi, G., T. Lundgren and L. Persson (2018). "The Rebound Effect in Swedish Heavy Industry." Energy Economics **71**: 140-148.
Amsler, C., A. Prokhorov and P. Schmidt (2016). "Endogeneity in stochastic frontier models." Journal of Econometrics **190**(2): 280-288.
Bardazzi, R., F. Oropallo and M. G. Pazienza (2015). "Do manufacturing firms react to energy prices? Evidence from Italy." Energy Economics **49**: 168-181.
Bartelsman, E. and W. Gray (1996). The NBER Manufacturing Productivit Data Base. Cambridge MA, National Bureau of Economic Research.
Bloom, N., C. Genakos, R. Martin and R. Sadun (2010). "Modern Management: Good for the Environment or Just Hot Air?*." The Economic Journal **120**(544): 551-572.
Bostian, M., R. Färe, S. Grosskopf and T. Lundgren (2016). "Environmental investment and firm performance: A network approach." Energy Economics **57**: 243-255.
Boyd, G., M. Doolin, E. Dutrow and S. Zhang (2017). "A New Benchmark of Energy Performance for Energy Management in U.S. and Canadian Integrated Steel Plants." Iron and Steel Technology(May): 53-61.
Boyd, G. and J. M. Lee (2016). Measuring Plant Level Energy Efficiency and Technical Change in the U.S. Metal-Based Durable Manufacturing Sector Using Stochastic Frontier Analysis, Center for Economic Studies, U.S. Census Bureau.
Boyd, G. and J. M. Lee (2017). Energy Efficiency, Technical Change and Price Responsiveness in Non-Energy Intensive Chemicals Manufacturing.
Boyd, G. and G. Zhang (2013). "Measuring improvement in energy efficiency of the US cement industry with the ENERGY STAR Energy Performance Indicator." Energy Efficiency **6**(1): 105-116.
Boyd, G. A. (2008). "Estimating Plant Level Manufacturing Energy Efficiency with Stochastic Frontier Regression." The Energy Journal **29**(2): 23-44.
Boyd, G. A. (2014). "Estimating the changes in the distribution of energy efficiency in the U.S. automobile assembly industry." Energy Economics **42**(0): 81-87.

Boyd, G. A. (2016). "Comparing the statistical distributions of energy efficiency in manufacturing: meta-analysis of 24 Case studies to develop industry-specific energy performance indicators (EPI)." Energy Efficiency: 1-22.
Boyd, G. A. and E. M. Curtis (2014). "Evidence of an “Energy-Management Gap” in U.S. manufacturing: Spillovers from firm management practices to energy efficiency." Journal of Environmental Economics and Management **68**(3): 463-479.
Boyd, G. A. and C. Delgado (2012). Measuring Improvement in the Energy Performance of the U.S. Corn Refining Industry. DUKE ENVIRONMENTAL ECONOMICS WORKING PAPER SERIES. Durham NC, NICHOLAS INSTITUTE FOR ENVIRONMENTAL POLICY SOLUTIONS**:** 1-18.
Boyd, G. A. and Y. Guo (2014). "An Energy Performance Indicator for integrated paper and paperboard mills: A new statistical model helps mills set energy efficiency targets." Paper360 **March/April**: 26-28.
Dalzell, N. M., G. A. Boyd and J. P. Reiter (2017). "Creating linked datasets for SME energy-assessment evidence-building: Results from the U.S. Industrial Assessment Center Program." Energy Policy **111**: 95-101.
Davis, S. J., C. Grim, J. Haltiwanger and M. Streitwieser (2012). "Electricity Unit Value Prices and Purchase Quantities: U.S. Manufacturing Plants, 1963–2000." Review of Economics and Statistics **95**(4): 1150-1165.
Energy Information Administration (1994). NEMS Industrial Module Documentation Report, U.S. Department of Energy.
Filippini, M. and L. C. Hunt (2011). "Energy Demand and Energy Efficiency in the OECD Countries: A Stochastic Demand Frontier Approach." The Energy Journal **32**(2): 59-80.
Filippini, M. and L. C. Hunt (2012). "US residential energy demand and energy efficiency: A stochastic demand frontier approach." Energy Economics **34**(5): 1484-1491.
Filippini, M. and L. C. Hunt (2015). "Measurement of energy efficiency based on economic foundations." Energy Economics **52, Supplement 1**: S5-S16.
Foster, L., C. Grim and J. Haltiwanger (2016). Reallocation In The Great Recession: Cleansing Or Not? Working Paper, Center for Economic Studies, U.S. Census Bureau.
Gerarden, T. D., R. G. Newell and R. N. Stavins (2017). "Assessing the Energy-Efficiency Gap." Journal of Economic Literature **55**(4): 1486-1525.
Gillingham, K. and K. Palmer (2014). "Bridging the Energy Efficiency Gap: Policy Insights from Economic Theory and Empirical Evidence." Review of Environmental Economics and Policy **8**(1): 18-38.
Greene, W. H. (2002). Fixed and Random Effects in Stochastic Frontier Models. New York Univesity Economics department working paper.
Huntington, H. (1995). "Been Top Down So Long It Looks Like Bottom Up." Energy.
Jaffe, A. B. and R. N. Stavins (1994). "The energy-efficiency gap What does it mean?" Energy Policy **22**(10): 804-810.
Jondrow, J., I. Materov, K. Lovell and P. Schmidt (1982). "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." Journal of Econometrics **19**(2/3): 233-238.
Kermeli, K., E. Worrell, W. Graus and M. Corsten (2017). Energy Efficiency and Cost Saving Opportunities for Ammonia and Nitrogenous Fertilizer Production An ENERGY STAR® Guide for Energy & Plant Managers. O. o. A. a. Radiation.
Kumbhakar, S. C., G. Lien and J. B. Hardaker (2014). "Technical efficiency in competing panel data models: a study of Norwegian grain farming." Journal of Productivity Analysis **41**(2): 321-337.
Lundgren, T., P.-O. Marklund and S. Zhang (2016). "Industrial energy demand and energy efficiency – Evidence from Sweden." Resource and Energy Economics **43**: 130-152.
Lundgren, T. and W. Zhou (2017). "Firm performance and the role of environmental management." Journal of Environmental Management **203**(Part 1): 330-341.

Lutz, B. J., P. Massier, K. Sommerfeld and A. Löschel (2017). Drivers of energy efficiency in German manufacturing: A firm-level stochastic

frontier analysis. ZEW Discussion Papers. Z. f. E. W. C. f. E. E. Research.
Neelis, M., E. Worrell and E. Masanet (2008). Energy Efficiency Improvement and Cost Saving Opportunities for the Petrochemical Industry An ENERGY STAR® Guide for Energy and Plant Managers, Ernest Orlando Lawrence Berkeley National Laboratory.
Nguyen, S. V. and M. L. Streitwieser (2008). "Capital-Energy Substitution Revisited: New Evidence from Micro Data." Journal of Economic and Social Measurement **33**(2-3): 129-153.
U.S. Energy Information Administration (2017). Annual Energy Outlook 2017 with projections to 2050. Washington DC.
United States Government Interagency Working Group on Social Cost of Carbon (2013). Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis - under Executive Order 12866. C. o. E. Q. Council of Economic Advisors, Department of Agriculture, Department of Commerce, Department of Energy, Department of Transportation, Environmental Protection Agency, National Economic Council, Office of Management and Budget, Office of Science and Technology and Department of the Treasury. Washington, DC.
US EPA (2015). Regulatory Impact Analysis for the Clean Power Plan Final Rule. O. o. A. Q. P. a. S. a. O. o. A. a. Radiation. Research Triangle Park, NC.
Zhang, S., T. Lundgren and W. Zhou (2016). "Energy efficiency in Swedish industry: A firm-level data envelopment analysis." Energy Economics **55**: 42-51.