Online Appendix to: Energy Cost Pass-Through in U.S. Manufacturing: Estimates and Implications for Carbon Taxes

Sharat Ganapati
Georgetown University

Joseph S. Shapiro
University of California, Berkeley
and NBER

Reed Walker
University of California, Berkeley
and NBER

A Recovering Marginal Costs, Output Elasticities, and Production Functions - For Online Publication

A.1 Recovering Marginal Costs

This section describes the methodology for calculating marginal costs, following Hall (1986); De Loecker and Warzynski (2012). Let Q_{it} denote the physical output Q of plant i in year t. Output is a function of variable inputs V_{it} (i.e., those not subject to adjustment costs) like materials and energy; dynamic inputs D_{it} like capital or sticky labor, which are subject to adjustment costs; and plant-specific productivity Ω_{it}: $Q_{it} = Q_{it}(V_{it}, D_{it}, \Omega_{it})$. We assume a firm minimizes the cost of the variable input(s), conditioning on the dynamic inputs. The firm solves the following Lagrangian:

$$\mathcal{L}(V_{it}, D_{it}, \lambda_{it}) = P_{it}^{V} V_{it} + R_{it} K_{it} + \lambda_{it}[Q_{it} - Q_{it}(V_{it}, D_{it}, \Omega_{it})]$$

Here P_{it}^{V} is the price of variable inputs, R_{it} is the price of dynamic inputs, and λ_{it} is the Lagrange multiplier.

The firm's first-order condition for a variable input like materials is

$$\frac{\partial \mathcal{L}}{\partial V_{it}} = P_{it}^{V} - \lambda_{it} \frac{\partial Q_{it}(\cdot)}{\partial V_{it}}$$

Rearranging terms for an optimum where $\partial \mathcal{L}/\partial V_{it} = 0$ and multiplying the right-hand side by $V_{it}Q_{it}/V_{it}Q_{it}$ and both sides by P_{it} shows how we recover markups:

$$\frac{P_{it}}{\lambda_{it}} = \left[\frac{\partial Q_{it}(\cdot)}{\partial V_{it}} \frac{V_{it}}{Q_{it}}\right] \left[\frac{P_{it}^{V} V_{it}}{P_{it} Q_{it}}\right]^{-1} \tag{1}$$

The left-hand side of this equation is the multiplicative markup μ_{it}, which equals prices divided the Lagrange multiplier. The Lagrange multiplier represents marginal costs, since it reflects the costs of relaxing the output constraint. The right-hand side is the product of two bracketed terms; we construct empirical analogues to both. The first is the elasticity of output with respect to a variable input, or the "output elasticity." We estimate the output elasticity from production functions, described in the next subsection. The second bracketed term is the cost of the variable input divided by the firm's revenue, or the "revenue share." Our data report the revenue share of each input.

We can then compute a time-varying, plant-level markup by using the estimated output elasticity of a variable input and the revenue share of that input. Since we observe plant-level unit prices, we then recover marginal costs from the accounting identity that price equals markups times marginal costs: $MC_{it} = P_{it}/\mu_{it}$ where MC_{it} is the marginal cost of plant i in year t, and μ_{it} is the (multiplicative) markup.

A.2 Recovering Output Elasticities from Production Functions

The previous subsection showed that estimating markups requires the output elasticity of a variable plant-level input like materials. We estimate this output elasticity by using proxy methods to estimate production functions (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves, and Frazer, 2015). We focus on production functions with a scalar, Hicks-neutral productivity term and estimate elasticities separately by industry, assuming common technology across firms and over time within an industry.[1] We show a Cobb-Douglas specification here to simplify exposition, though our results use a more flexible translog, gross-output production function:[2]

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it} \tag{2}$$

Throughout the paper, lowercase represents variables in logs. Here y_{it} represents a plant's output quantity. We use output quantity rather than revenues here to avoid well-known bias in revenue-based productivity estimates.[3] Firms use three inputs: capital, labor, and materials (k_{it}, l_{it}, and m_{it}). Materials includes energy inputs in addition to other intermediate inputs used for production. The parameter vector which we estimate, $\beta \equiv (\beta_k, \beta_l, \beta_m)$, contains the output elasticities of these three inputs. The term ω_{it} represents productivity, which is known to the firm when making static input decisions but unobserved to the econometrician. The residual ϵ_{it} includes measurement error and unanticipated shocks to output.

Ordinary least squares estimates of equation (2) may suffer from omitted variables bias due to the unobserved productivity term ω_{it} (Marschak and Andrews, 1944). A firm observes its productivity, so input choices k_{it}, l_{it}, and m_{it} may depend on it, but productivity directly affects output, and data do not report it.

To address the possible omitted variable bias associated with OLS estimates of equation (2), we use control-function or proxy methods to control for the unobserved and omitted productivity term. Consider a general demand function for materials:[4] $m_{it} = m_t(k_{it}, l_{it}, \omega_{it})$.

[1] Output elasticities could in principle differ by industry and time period. Such flexible output elasticities are difficult to estimate with our data, however, since we have few years of data, require one lag to construct instruments, and have few observations for most industries.

[2] Translog coefficients are the same across firms within an industry. Markups and output elasticities, however, differ across firms within an industry, because input demands differ across firms. This is an advantage of translog over Cobb-Douglas production functions, which would have the same output elasticity across firms within an industry. Under Cobb-Douglas, all variation in markups across firms within an industry would come from revenue shares.

[3] Output quantity addresses the distinction between revenue and physical total factor productivity (Foster, Haltiwanger, and Syverson, 2008). Unobserved variation in input prices may also bias production function coefficients (De Loecker, Goldberg, Khandelwal, and Pavcnik, 2016). The homogeneity of our products potentially gives less scope for input price variation and associated bias (De Loecker and Goldberg, 2014). We have explored specifications that attempt to control for any remaining input price variation using a polynomial in the output price, and results are largely similar. Our dataset reports expenditures on inputs though not input quantities; we are unaware of any production function estimates in any setting using input quantity data for all inputs rather than price-deflated input expenditures.

[4] We focus on materials as a variable input into the production function, where materials include both purchased intermediates as well as energy input expenditures. In theory, we could estimate a separate output elasticity for energy in the production function. In practice, adding a fourth input into a translog production function substantially increases the number of parameters to be estimated. With relatively small sample sizes and relatively few degrees of freedom, output elasticities become more sensitive and less robust. Moreover, many indirect energy input costs are embodied in material expenditures through feedstock purchases for example.

Assuming that $m_t(\cdot)$ is strictly monotonic in inputs, we invert this input demand equation to solve for productivity as a function of the observable inputs:

$$\omega_{it} = m_t^{-1}(m_{it}, k_{it}, l_{it})$$

This inversion provides a control function for productivity.[5]

We apply this approach in two steps, following Ackerberg, Caves, and Frazer (2015). The first step regresses plant output y_{it} on a function $\phi_t(\cdot)$ of observed inputs. This first step is designed to purge output data of measurement error and unanticipated shocks to output ϵ_{it}:

$$y_{it} = \phi_t(k_{it}, l_{it}, m_{it}) + \epsilon_{it}$$

We approximate $\phi_t(\cdot)$ using a polynomial expansion. We use estimates from this first step to calculate ϵ_{it} from

$$\hat{\epsilon}_{it} = y_{it} - \hat{\phi}_t(k_{it}, l_{it}, m_{it}) \tag{3}$$

where $\hat{\phi}$ contains the fitted values from this first step, and $\hat{\epsilon}_{it}$ are the residuals from this regression. Since ϵ_{it} contains measurement error and unanticipated shocks to production, we can use it to obtain a measure of output which is purged of both. After this first step, the only missing information needed to know the output elasticity vector β is the productivity vector ω_{it}. Given any candidate elasticity vector $\tilde{\beta}$, we can estimate productivity by manipulating equations (2) and (3) to get

$$\omega_{it}(\tilde{\beta}) = \widehat{\phi}_{it} - \tilde{\beta}_k k_{it} - \tilde{\beta}_l l_{it} - \tilde{\beta}_m m_{it} \tag{4}$$

The second step selects the coefficient vector that best fits the data by relying on the law of motion for productivity. We follow Ackerberg, Caves, and Frazer (2015) and assume that productivity follows a first-order Markov process.[6] We define productivity shocks ξ_{it} as the difference between productivity and the expectation of last period's productivity given last period's information set $\mathbb{I}_{it-1}$:

$$\xi_{it} = \omega_{it} - \mathbb{E}[\omega_{it}|\mathbb{I}_{it-1}]$$

[5]Inverting materials demand to recover productivity requires a one-to-one mapping between plant-level productivity and materials. This assumption fails if unobserved plant-level variables besides productivity drive changes in materials or if there is measurement error in materials. Alternative production function estimators, such as the dynamic panel methods developed by Blundell and Bond (2000) are not appropriate in our setting since we have few time periods to construct differences and lags. Some evidence suggests these may not be first-order concerns. Syverson (2004) finds robustness among producer TFP measures (and hence output elasticities) for one of our industries, ready-mixed concrete, with a specification incorporating idiosyncratic demand shocks. Van Biesebroeck (2007) also finds high TFP correlations across various measurement alternatives. Given the strong assumptions needed to estimate output elasticities, however, subsequent sections explore alternative methods to characterize incidence in the absence of information on output elasticities or markups and marginal costs.

[6]We use the AR(1) process to derive a plausibly exogenous productivity shock ξ_{it} along the lines of Ackerberg, Caves, and Frazer (2015). We have also allowed for the potential of additional lagged decision variables to affect current productivity outcomes (in expectation) in order to accommodate the concerns raised by De Loecker (2011) pertaining to the exogeneity of productivity process. For example, we have allowed productivity to depend on export status and the nonrandom exit of firms (De Loecker, 2011; Olley and Pakes, 1996). In practice, our output elasticity estimates are not particularly sensitive to these modifications.

where $\mathbb{E}$ is the expectation operator. Equivalently, ξ_{it} represents the component of current productivity which was unexpected at time $t-1$.

The second step estimates the production function coefficients using the assumption that this productivity innovation must be orthogonal to a set of current and lagged input demands d_{it}. We summarize these conditions as

$$\mathbb{E}[\xi_{it}(\beta)d_{it}] = 0 \tag{5}$$

With the translog production function we use for the empirical implementation, the vector d_{it} is

$$d_{it} = \{l_{it}, m_{it-1}, k_{it}, l_{it}^2, m_{it-1}^2, k_{it}^2, l_{it}m_{it-1}, l_{it}k_{it}, m_{it-1}k_{it}\}$$

These moments above are similar to those suggested by Ackerberg, Caves, and Frazer (2015). They exploit the fact that capital and labor have adjustment costs, and that lagged capital and labor should not be correlated with the current productivity innovation. We use lagged rather than current materials to identify the materials coefficients since current material expenditures may react to contemporaneous productivity innovations. For lagged materials to be a valid instrument for current materials, input prices must to be correlated over time.

Finally, we use generalized method of moments to choose the production function coefficients β which minimize the moment conditions in equation (5). With translog production functions, the coefficients β combined with input data give the output elasticities $\hat{\theta}$:[7]

$$\hat{\theta}_{it} = \theta(\hat{\beta}, l_{it}, m_{it}, k_{it})$$

These output elasticities vary by plant and by year. This leads to plant-year variation in markups (i.e., different markups for each plant within an industry-year) that is driven by both changes in the plant-level revenue share for the variable input (e.g. materials) and changes in the mix of inputs used for production.

B Appendix Figures and Tables

[7] The estimated output elasticity for materials, for example, is $\hat{\theta}_{it} = \hat{\beta}_m + 2\hat{\beta}_{mm}m_{it} + \hat{\beta}_{lm}l_{it} + \hat{\beta}_{km}k_{it} + \hat{\beta}_{lmk}l_{it}k_{it}$.

Table B1: Pass-Through Rate of Marginal Costs into Unit Prices, by Product: Fuel-Mix Instrumental Variables

	(1) Boxes	(2) Bread	(3) Cement	(4) Concrete	(5) Gasoline	(6) Plywood
	Fuel Shift-Share Instrument (with Year FE)					
Marginal Costs	0.946***	0.214	0.706***	0.743***	0.491***	0.825***
	(0.031)	(0.443)	(0.067)	(0.098)	(0.139)	(0.076)
N	1414	248	229	3369	284	139
First Stage F-Statistic	12.60	4.25	18.50	15.22	4.73	77.18
Year FE	X	X	X	X	X	X
Plant FE	X	X	X	X	X	X
State-Trends FE	X	X	X	X	X	X

Notes: This table presents regression coefficients from 7 separate regressions; one per column. Each column represents a separate sample, where the sample is indicated in the column headings. An observation is a plant-year. The dependent variable is the plant-level unit-price, and the independent variable is plant-level marginal cost. Marginal cost is instrumented by the interactions between national fuel prices for industrial production and 2-year lagged industry energy expenditure shares. All regressions include plant fixed effects, year fixed effects, and state-specific trends. Standard errors are in parentheses and are clustered by state. Regressions are weighted by Census sampling weights. ***,**,* denotes statistical significance at the 1, 5, and 10 percent levels, respectively. See text for details. Source: Census and Annual Survey of Manufacturers, MECS, EIA-SEDS.

Table B2: Pass-Through Rate of Marginal Costs into Unit Prices, in Levels, by Product: Instrumental Variables

	(1) Boxes	(2) Bread	(3) Cement	(4) Concrete	(5) Gasoline	(6) Plywood
	Panel A: Baseline - Electricity Price Instrument					
Marginal Costs	1.581***	0.029	1.585***	0.759***	0.461***	0.625***
	(0.052)	(0.157)	(0.219)	(0.061)	(0.132)	(0.225)
N	1414	308	293	3369	345	163
First Stage F-Statistic	44.59	4.00	80.03	12.49	7.15	17.04
Plant FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
State-Trends FE	X	X	X	X	X	X
	Panel B: Region-Year FE - Electricity Price Instrument					
Marginal Costs	1.597***	0.019	1.631***	0.717***	0.302**	0.755***
	(0.111)	(0.096)	(0.192)	(0.064)	(0.123)	(0.160)
N	1414	308	293	3369	345	163
First Stage F-Statistic	2.83	3.87	3.64	12.66	8.99	233.0
Plant FE	X	X	X	X	X	X
Region×Year FE	X	X	X	X	X	X
State-Trends FE	X	X	X	X	X	X

Notes: This table presents regression coefficients from 14 separate regressions; one per column in each of the two panels. Each column represents a separate sample, where the sample is indicated in the column headings. An observation is a plant-year. The dependent variable is the plant-level unit-price, and the independent variable is plant-level marginal cost. Marginal cost is instrumented by the interactions between national fuel prices for electricity generation and 5-year lagged electricity generation shares. Standard errors are in parentheses and are clustered by state. Regressions are weighted by Census sampling weights. ***,**,* denotes statistical significance at the 1, 5, and 10 percent levels, respectively. See text for details. Source: Census and Annual Survey of Manufacturers, MECS, EIA-SEDS.

Table B3: Plant-Level Incidence: Mean and Standard Deviation

	(1) Incidence Mean	(2) Incidence Standard Deviation	(3) Coefficient of Variation
Boxes	0.664	0.032	0.05
Bread	0.495	0.119	0.24
Cement	0.506	0.044	0.09
Concrete	0.605	0.092	0.15
Gasoline	0.364	0.034	0.09
Plywood	0.644	0.074	0.11

Notes: This table presents additional summary information pertaining to within-industry incidence heterogeneity. Column (1) presents the mean incidence for a given industry, where the mean is taken over plant-level incidence measures. Plant-level incidence measures were created using plant-level markup estimates, combined with plant-level pass-through rates and industry-level demand elasticities. Column (1) reflects the unweighted mean over these plant-level observations. Column (2) represents the standard deviation of these incidence estimates within an industry. Column (3) represents the coefficient of variation, which is defined as the ratio of the standard deviation of a sample relative to it's mean. Thus, Column (3) is created by dividing Column (2) by Column (1).

Table B4: Demand Elasticity Estimates

	(1) Boxes	(2) Bread	(3) Cement	(4) Concrete	(5) Gas	(6) Plywood
	Panel A: OLS					
Demand Elasticity (ϵ_D)	-0.377**	-0.273	-0.387	-0.657	-0.0454	0.00469
	(0.121)	(0.211)	(0.286)	(0.505)	(0.0748)	(0.196)
	Panel B: Productivity IV Estimates					
Demand Elasticity (ϵ_D)	-2.762**	-5.233	-2.902**	-4.275*	-0.131	-1.926*
	(0.894)	(9.187)	(1.054)	(1.980)	(0.111)	(0.820)
N	100	25	25	25	25	50
First Stage F-Statistic	11.71	0.267	9.312	5.209	8.673	8.181
Year Trend	X	X	X	X	X	X

Notes: This table presents 12 separate regressions, 6 per panel. An observation is the yearly change in an industry-year, where the dependent variable in all regressions is Δlog(quantity). The independent variable is Δlog(output price). Panel A presents OLS estimates, separately by industry. Panel B presents estimates where price is instrumented with changes in industry level total factor productivity. Total factor productivity is constructed using a quantity-based productivity index. The index is constructed by subtracting log inputs from log outputs using industry-level cost shares as proxies for output elasticities. We use capital, materials, labor, and energy inputs, where capital, materials, and energy are deflated by industry-year input price deflators, and labor is measured in production hours. Standard errors are computed using Newey and West (1987). Source: NBER-CES Manufacturing Database.

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83(6), 2411–2451.

Blundell, R., and S. Bond (2000): "GMM estimation with persistent panel data: an application to production functions," *Econometric reviews*, 19(3), 321–340.

De Loecker, J. (2011): "Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity," *Econometrica*, 79(5), 1407–1451.

De Loecker, J., and P. K. Goldberg (2014): "Firm Performance in a Global Market," *Annu. Rev. Econ.*, 6(1), 201–227.

De Loecker, J., P. K. Goldberg, A. Khandelwal, and N. Pavcnik (2016): "Prices, Markups and Trade Reform," *Econometrica*, 84(2), 445–510.

De Loecker, J., and F. Warzynski (2012): "Markups and Firm-Level Export Status," *The American Economic Review*, 102(6), 2437–2471.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Hall, R. E. (1986): "Market structure and macroeconomic fluctuations," *Brookings papers on economic activity*, pp. 285–338.

Levinsohn, J., and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.

Marschak, J., and W. H. Andrews (1944): "Random simultaneous equations and the theory of production," *Econometrica, Journal of the Econometric Society*, pp. 143–205.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.

Syverson, C. (2004): "Market Structure and Productivity: A Concrete Example," *Journal of Political Economy*, 112(6), 1181–1222.

Van Biesebroeck, J. (2007): "Robustness of Productivity Estimates," *The Journal of Industrial Economics*, 55(3), 529–569.