Convex Supply Curves*

Christoph E. Boehm[1] and Nitya Pandalai-Nayar[1]

[1]UT Austin

March 2019

Abstract

We provide evidence that industries' supply curves are convex. To guide our empirical analysis we develop a putty-clay model in which the capacity utilization rate is a sufficient statistic for the slope of the supply curve. Using data on capacity utilization, we estimate the supply curve using three different instruments—foreign demand shocks, demand shocks from downstream industries, and exchange rate shocks—and find robust evidence for convexity. Supply curves are essentially flat at low levels of capacity utilization but increasing at higher levels. Further, industries with low initial capacity utilization rates expand production more after demand shocks than industries that are producing close to their capacity limit. The nonlinearity we identify has implications for a number of applications in macro and international economics, including that responses to shocks are state-dependent and that the Phillips curve is convex.

JEL Codes: F14, F16, F23

Keywords: Slack, Utilization, Capacity, Supply Curves, State-Dependence, Fiscal Policy, Monetary Policy, Exchange Rate Shocks

*chris.e.boehm@gmail.com and npnayar@utexas.edu. We thank our discussants Yuriy Gorodnichenko and Hikaru Saijo. The paper also benefited from conversations with Mark Aguiar, Pol Antras, Saroj Bhattarai, Oli Coibion, Javier Cravino, Aaron Flaaen, Oleg Itskhoki, Andrei Levchenko, Chris Sims, and Jaume Ventura as well as with seminar participants a number of different locations. Pandalai-Nayar thanks the ECB for the Baron Alexandre Lamfalussy Fellowship which supported this research. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau.

1 Introduction

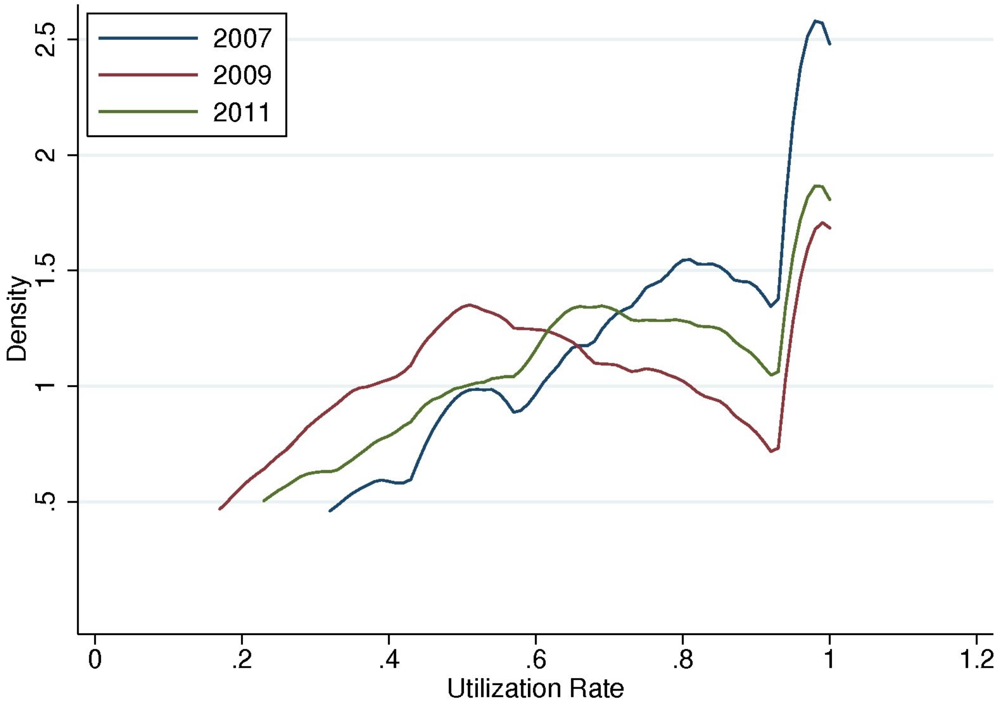
Up until a decade ago, most commonly used frameworks in macroeconomics were either explicitly linearized, or almost linear when solved exactly (Parker, 2011). A limitation of such frameworks is that they can lead to imprecise conclusions on aspects of the economy that are fundamentally nonlinear. While recent research has explicitly emphasized the role of such nonlinearities (e.g., Kaplan and Violante (2014), Baqaee and Farhi (2017), among many others), disagreement in particular on the empirical evidence persists. For instance, there is a lack of consensus as to whether the fiscal multiplier varies with the state of the business cycle. While Auerbach and Gorodnichenko (2012; 2013a; 2013b) argue that the multiplier is likely large during severe downturns, Ramey and Zubairy (2018) find that such state dependence is small or nonexistent. In other applications, the empirical evidence is largely lacking. Baqaee and Farhi (2017) highlight that disregarding nonlinearities in models with input-output linkages can have large quantitative implications. As of now, however, there is limited direct evidence on how important nonlinearities in such settings are.[1]

In this paper we document robust evidence for one particular nonlinearity, namely that industry's supply curves are convex. We show that these supply curves are essentially flat at low levels of production, but increasing at higher levels—in particular, when production is close to capacity. Such curvature implies that industries' production responses to shocks depend on the initial degree of capacity utilization. Compared to an initial capacity utilization rate at the 95th percentile, the response to the same generic demand shock is approximately twice as large if the initial utilization rate is instead at the 5th percentile. Convex supply curves can potentially explain a number of additional phenomena in macro and international economics. Examples include the missing disinflation which many advanced economies experienced during the Great Recession, a distribution of GDP growth rates that is skewed to the left, and so forth.

We begin our analysis with the following three observations. First, as Figure 1 illustrates, a significant fraction of U.S. manufacturing establishments produces at "full capacity" as defined in the Survey of Plant Capacity, or equivalently, at a utilization rate of one. These plants presumably have limited room for increasing production in the short run. The remaining plants produce below their reported capacity, at times far below: Utilization rates between 0.2 and 0.5 are not uncommon. Second, the fraction of plants with utilization rates of unity are procyclical. In 2007 a large fraction of plants produced at full capacity and the

[1] See Huo, Levchenko, and Pandalai-Nayar (2019) for some evidence.

Figure 1: Densities of plant capacity utilization



Notes: The data are from the QSPC of the U.S. Census Bureau. The figure shows kernel density estimates which are truncated below the 5th and above the 95th percentile due to Census disclosure requirements.

density displays a mode at around 0.8. By 2009 the distribution has shifted to the left and has a modal point at approximately 0.5. The 2011 density reflects partial recovery relative to 2009 but utilization rates are still below those of 2007. Third, plants produce below their available capacity predominantly because they are not able to sell their products. For the time period from 2013q1 to 2018q2 for which public data is available, 78.4 percent of plant managers cites insufficient orders as the main reason for producing below capacity. The second most cited option is chosen by 11.5 percent of respondents (insufficient supply of local labor force/skills). In summary, these three observations suggest that firms first build up capacity and subsequently produce and sell subject to a capacity limit. We discuss details regarding the data and measurement as well as further facts in Appendix A.

To guide our empirical analysis we develop a simple putty-clay framework that features this notion of capacity utilization. Drawing on earlier work by Fagnart, Licandro, and Portier (1999), firms invest into a set of factors that are fixed in the short run and, once chosen, determine the firm's maximum productive capacity. When the demand for a firm's good materializes sufficiently high, production becomes constrained by capacity. These constrained firms are locally unresponsive to shocks because any changes in demand are absorbed in changes in the markup. The framework permits simple aggregation to the industry level, where it generates a supply curve that is convex in capacity utilization. A key feature of the model is that the capacity utilization rate is a sufficient statistic for the slope of the supply curve. While we use the model to motivate our main empirical specifications, we note that as long as our identification assumptions hold the results do not depend on the particular framework adopted here.

Our main contribution is to provide empirical evidence that industry's supply curves are convex. We estimate both the structural form of the supply curve and the reduced form using three alternative instruments to trace out the slope and curvature of the supply curve. First, we use a version of the World Import Demand (WID) instrument (Hummels et al., 2014). This instrument assumes that appropriately purified changes in foreign import demand are uncorrelated with the industry's unobserved supply shocks. Second, and building on Shea (1993a,b) we construct an instrument from changes in downstream demand. The idea of this instrument is to alleviate simultaneity concerns in production networks by isolating variation from unidirectional linkages. Third, we consider changes in industry's effective exchange rates. Conditional on holding industry's costs constant, depreciations in the exchange rate stimulate demand from abroad.[2] We emphasize that our estimates are comparable for all three instruments.

The estimates suggest that supply curves are highly elastic at low levels of capacity utilization. At capacity utilization rates below the 5th percentile, we cannot reject the null hypothesis that industries' supply curves are horizontal. This contrasts to an estimated inverse supply elasticity of approximately 0.3 at the median and in excess of 0.4 above the 95th percentile. We also find that the production response to the same sized demand shock is decreasing in the capacity utilization rate and approximately twice as large at the 5th percentile when compared to the 95th percentile. Our estimation uses the measures of capacity utilization from the Federal Reserve Board (FRB) which are close empirical analogues to the corresponding object in the model.

Our findings are relevant for a number of applications in macro and international economics. First, our estimates imply that responses to shocks are state-dependent. Policies

[2] We have also tried an additional defense spending instrument, but it is too weak at the industry level to deliver useful results.

which stimulate demand will raise output more when implemented in times of slack. These results are consistent with the findings of Auerbach and Gorodnichenko (2012, 2013a,b), but are not limited to government spending shocks. Further applications—to be completed.

Our empirical strategy uses capacity utilization as a sufficient statistic for estimation. Capacity utilization measures how much a firm or plant produces relative to how much it can produce (capacity). It is important to note that this concept is different from, though related to, *capital utilization* which measures the fraction of time the capital stock operates. A large literature in macroeconomics has studied models in which capital services vary at high frequencies due to a utilization choice (see, e.g. Greenwood, Hercowitz, and Huffman, 1988, Bils and Cho, 1994, Cooley, Hansen, and Prescott, 1995, Burnside and Eichenbaum, 1996, Fagnart, Licandro, and Sneessens, 1997, Gilchrist and Williams, 2000, Hansen and Prescott, 2005).[3] Relative to that literature the nonlinear implications of capacity utilization have received less attention. Two exceptions are Michaillat (2014), who develops the idea that slack in the labor market (or a convex labor supply curve) leads to state-dependent fiscal multipliers, and Kuhn and George (2017) who study the quantitative implications of convex supply curves in general equilibrium.[4] Unlike models in which utilization is subject to convex costs, our putty-clay framework allows for an explicit distinction between the short run and the long run at which capacity adjusts.

Our paper also complements earlier work on state-dependent responses to shocks. These papers include, but are not limited to, Weise (1999), Peersman and Smets (2005), Lo and Piger (2005), Baum, Poplawski-Ribeiro, and Weber (2012), Owyang, Ramey, and Zubairy (2013), Fazzari, Morley, and Panovska (2015), Jordà, Schularick, and Taylor (2017), and those cited above. All of these papers study one particular shock, rather than a mechanism—as we do. Further, we conduct our empirical analysis at the industry level to obtain more statistical power.

We begin in Section 2 with presenting a simple model that features utilization of capacity and motivates our empirical strategy. After discussing the data and identification we present our baseline empirical results in Section 3. We discuss the implications of our findings for a number of applications and provide guidance on how to calibrate convex supply curves in Section 4. Section 5 concludes.

[3] See also Alvarez-Lois (2004, 2006).

[4] For empirical work on capacity and capital utilization, see Shapiro (1989), Stock and Watson (1999), and Gorodnichenko and Shapiro (2011).

2 Theoretical framework

This section lays out the theoretical foundation for estimating supply curves. We begin with a simple and stylized putty-clay model which features a concept of capacity that aligns well with measured capacity in the data. When aggregated to the industry level, this framework generates a supply curve which is typically increasing and convex in capacity utilization. We subsequently specify the demand side and present our estimating equations.

2.1 A simple theory of capacity constraints and convex supply curves

Our framework features a competitive aggregating firm and monopolistically competitive intermediate goods firms. In order to generate a notion of capacity and utilization we assume a putty-clay-type production function (as in Fagnart, Licandro, and Portier, 1999) which requires firms to choose their maximum scale prior to making the actual production decision. If demand materializes sufficiently high, production will be constrained by capacity.

2.1.1 Aggregating firm

A competitive aggregating firm uses a unit continuum of varieties, indexed j, as inputs into a constant elasticity of substitution (CES) aggregator to produce the industry's composite good,

$$X_t = \left(\int_0^1 \omega_t (j)^{\frac{1}{\theta}} x_t (j)^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}}. \tag{1}$$

Parameter θ is the elasticity of substitution and the weights ω represent firm-specific and time-varying demand shocks for intermediate goods producers. For simplicity, we assume these shocks are drawn independently and identically from distribution G with $\mathbb{E}[\omega] = 1$ and $\mathbb{E}[\omega^3] < \infty$.

Taking prices as given, the final goods firm maximizes profits subject to the production function (1). The resulting input demand curves are

$$x_t (j) = \omega_t (j) X_t \left(\frac{p_t^x (j)}{P_t^X} \right)^{-\theta} \tag{2}$$

for all j, where the industry's price index is given by

$$P_t^X = \left(\int_0^1 \omega_t (j) p_t^x (j)^{1-\theta} dj \right)^{\frac{1}{1-\theta}}. \tag{3}$$

2.1.2 Intermediate goods producers

Consistent with our motivating facts in Section 1, we assume that a firm's capacity can limit production in the short run. Following Fagnart, Licandro, and Portier (1999), the firm has to decide ex-ante on the maximum of variable inputs, $\bar{v}_t$, that it can employ (or process) in the short run. Since short-run variable inputs v_t include primarily production workers and intermediates, $\bar{v}_t$ has a natural interpretation as the number of workstations or the capacity to process intermediates. To preserve clarity we drop the index j throughout this section.

Production and capacity Let q_t denote the firm's *idiosyncratic production capacity* which is predetermined within the period. The firm's production function is

$$x_t = q_t \frac{v_t}{\bar{v}_t}, \quad \text{where} \quad v_t \leq \bar{v}_t. \tag{4}$$

That is, production x_t is linear in short-run variable inputs v_t, but subject to an upper bound in the short run because the variable inputs v_t cannot exceed the predetermined value $\bar{v}_t$.

Letting z_t denote productivity and k_t capital, firm's production capacity takes the form

$$q_t = z_t F\left(k_t, \bar{v}_t\right). \tag{5}$$

The function F is increasing in k_t and $\bar{v}_t$, and exhibits constant returns to scale in its two arguments. The latter assumption implies that firm's actual production can be written as $x_t = z_t F\left(\kappa_t, 1\right) v_t$ where $\kappa_t = k_t / \bar{v}_t$. That is, the marginal product of v_t is $z_t F\left(\kappa_t, 1\right)$ which is increasing in z_t and κ_t. Letting p_t^v denote the price of the variable input bundle v_t, short-run marginal costs are

$$mc_t = \frac{p_t^v}{z_t F\left(\kappa_t, 1\right)}. \tag{6}$$

Dynamic problem Firms own their capital stock k and maximize the present value of profits. We allow firm's investment to be subject to possibly non-convex adjustment costs $\phi\left(i, k\right)$. The firm's Bellman equation is then

$$V\left(k, \bar{v}, z, \omega\right) = \max_{x,i,v,\bar{v}'} \left\{ p^x x - p^v v - p^i i - \phi\left(i, k\right) + \frac{1}{1+r} \mathbb{E}\left[V\left(k', \bar{v}', z', \omega'\right)\right] \right\},$$

where the maximization is subject to

$$x \leq q, \tag{7}$$

$$k' = (1-\delta)\,k + i, \tag{8}$$

as well as equations (2), (4), and (5). Equation (7) is the capacity constraint and (8) is the standard capital accumulation equation. For simplicity, we assume that productivity z only has an industry-specific and an aggregate (i.e. economy wide), but no firm-specific component.

Our estimation strategy does not require us to take a stance on the functional form of adjustment costs ϕ. Nor are the precise features of the firm's investment decision or the choice of $\bar{v}'$ important for the estimation. What matters for our estimation is the evolution of capacity, or more precisely the *industry's* capacity—which we directly observe in the data. We discuss the role of changes in capacity for the estimation below and relegate a discussion the firm's choices of k' and $\bar{v}'$ to Appendix B.

Price setting If the firm operates below its capacity limit, it sets prices at a constant markup over marginal costs. Once production is constrained by capacity, however, the firm raises its markup so as to equate the quantity demanded to its production capacity. Formally,

$$p^x = \frac{\theta}{\theta-1}\left(mc+\psi\right), \qquad \psi = 0 \quad \text{whenever} \quad x < q,$$

where ψ is the multiplier on the capacity constraint (7). In this baseline version of the model rising markups are the key mechanism generating the convex supply curve. In Appendix B we discuss a number of alternative mechanisms that lead to such convexity. These include rationing (in the presence of sticky prices) and kinks in the cost function (for instance due to a shift premium).

Since the idiosyncratic demand shock ω is the only source of heterogeneity, there exists a threshold variety $\bar{\omega}$ above which firms' production is constrained by their capacity. A lower value of $\bar{\omega}$ implies that more firms are capacity constrained. We next show that $\bar{\omega}$ plays a critical role for characterizing the degree to which the industry uses its productive capacity.

2.1.3 Industry capacity and utilization

Using equation (1) the industry's output can be written as

$$X\left(q_t,\bar{\omega}_t\right) = q_t\left(\left(\bar{\omega}_t\right)^{-\frac{\theta-1}{\theta}}\int_0^{\bar{\omega}_t}\omega dG\left(\omega\right) + \int_{\bar{\omega}_t}^{\infty}\left(\omega\right)^{\frac{1}{\theta}}dG\left(\omega\right)\right)^{\frac{\theta}{\theta-1}}.$$

In particular, the industry's output is only a function of the common idiosyncratic plant capacity q_t, and the threshold variety $\bar{\omega}_t$.

We define the *industry's capacity* as the level of output that would be attainable if every intermediate firm produced at full capacity, that is

$$Q(q_t) := \lim_{\bar{\omega}_t \to 0} X(q_t, \bar{\omega}_t).$$

Further, we define the industry's utilization rate as the ratio of actual production to full capacity production,

$$u_t := \frac{X(q_t, \bar{\omega}_t)}{Q(q_t)}. \tag{9}$$

Note that this definition aligns well with its empirical counterpart. The Federal Reserve measures capacity utilization at the industry level by dividing an index of industrial production, i.e. a measure of gross output, by an estimate of capacity.

Lemma 1. *The utilization rate as defined in (9) has the following properties:*

1. $u_t \in [0,1]$ *is only a function of* $\bar{\omega}_t$*:* $u_t = u(\bar{\omega}_t)$
2. $\lim_{\bar{\omega} \to 0} u(\bar{\omega}) = 1$*,* $\lim_{\bar{\omega} \to \infty} u(\bar{\omega}) = 0$
3. $u' < 0$
4. *The sign of* u'' *is ambiguous*

The lemma highlights that the industry's utilization rate is only a function of the threshold value $\bar{\omega}_t$ above which firms produce at full capacity. The utilization rate approaches zero if no firm produces at full capacity and it tends to one if all firms become capacity constrained. Further, u is decreasing everywhere, and thus u is invertible and we can write $\bar{\omega}_t = \bar{\omega}(u_t)$. We will make extensive use of this property, both for the remainder of the theoretical analysis and when taking the model to the data.

2.1.4 The supply curve

One immediate application of the invertibility of u is that the industry's price index (3) can be written as

$$\ln P_t^X = \mathcal{M}(\ln u_t) + \ln(mc_t). \tag{10}$$

This (inverse) supply curve is the starting point for our empirical analysis. It depends on the industry's marginal costs mc_t, and the industry's average markup $\mathcal{M}$. This markup is only a function of the industry's utilization rate, or equivalently, of output relative to capacity. For the subsequent empirical analysis it is convenient to define the markup as a function of the *logarithm* of the utilization rate.

Proposition 1. *$\mathcal{M}$ has the following properties:*

1. $\mathcal{M}' \geq 0$
2. $\lim_{u \to 0} \mathcal{M}(\ln u) = \ln \frac{\theta}{\theta - 1}$, $\lim_{u \to 1} \mathcal{M}(\ln u) = \infty$
3. $\lim_{u \to 0} \mathcal{M}'(\ln u) = 0$, $\lim_{u \to 1} \mathcal{M}'(\ln u) = \infty$
4. *Without further restrictions on G, the sign of $\mathcal{M}''$ is generally ambiguous.*

Because $\mathcal{M}$ is increasing in utilization everywhere, the industry's supply curve (10) is upward-sloping. In contrast to standard models, the industry's markup rises when production X_t increases *relative* to capacity Q_t. As utilization rises, more suppliers become capacity constrained and those that are constrained respond by raising their markups. As the utilization rate approaches one, all suppliers become constrained and $\mathcal{M}$ and its derivative tend to infinity. Conversely, when the utilization rate tends to zero, fewer and fewer suppliers are capacity constrained. As a result $\mathcal{M}$ tends to $\ln \frac{\theta}{\theta-1}$ and its derivative to zero. While $\mathcal{M}(\ln u)$ is convex everywhere for many choices of G, it is possible to construct examples in which it is locally concave. Thus, whether $\mathcal{M}$ is convex in the relevant range of utilization remains an empirical question which we will address below. (Add figure here on how the supply curve depends on θ and G.)

2.2 Demand and market clearing

We next specify the demand side of the model and impose market clearing. As we demonstrate below, the estimation of the reduced form benefits from a detailed specification of the demand side which closely resembles industry's sales patterns in the data.

Each industry sells its product both domestically and abroad. For domestic sales we distinguish sales to downstream industries in the form of intermediates and final sales of consumption and investment goods. We assume for simplicity that demand (locally) takes the constant elasticity form.

Domestic final demand Depending on whether industry i produces a consumption or investment good, domestic final demand takes the form

$$\begin{aligned} C_{i,t} &= \omega_{i,t}^C C_t \left(\frac{P_{i,t}}{P_t^C}\right)^{-\sigma}, \\ I_{i,j,t} &= \omega_{i,j,t}^I I_{j,t} \left(\frac{P_{i,t}}{P_{j,t}^I}\right)^{-\sigma}. \end{aligned}$$

Here, C_t are real personal consumption expenditures (PCE), and P_t^C is the PCE price index. Similarly, $I_{j,t}$ is real investment into goods of category j (e.g. equipment investment), and $P_{j,t}^I$ is the associated price index. The elasticity σ parameterizes the substitutability of varieties within each of these aggregates. Unlike the quantity C_t and price P_t^C, $\omega_{i,t}^C$ is an *unobserved* demand shifter (and analogously $\omega_{i,j,t}^I$).

Intermediate demand Industry i further sells its output to other industries downstream. Letting $M_{j,t}$ denote the aggregate of industry j's purchases of intermediates, and $P_{j,t}^M$ the corresponding price index, its demand for industry i's output is

$$M_{i,j,t} = \omega_{i,j,t}^M M_{j,t} \left(\frac{P_{i,t}}{P_{j,t}^M}\right)^{-\sigma}.$$

Again, $\omega_{i,j,t}^M$ is an unobserved demand shock.

Foreign demand Exports abroad constitute an additional component of industry i's demand. Analogous to the earlier specification, we assume that demand of destination d is given by

$$EX_{i,d,t} = \omega_{i,d,t}^{\text{ex}} EX_{d,t} \left(\frac{P_{i,d,t}^*}{P_{d,t}^{\text{ex},*}}\right)^{-\sigma}.$$

Here, prices with asterisks are measured in foreign currency units. The dollar-denominated price for sales abroad is $P_{i,t}^X = \mathcal{E}_{d,t} P_{i,d,t}^{X,*}$, where $\mathcal{E}_{d,t}$ is the nominal exchange rate in U.S. dollars per unit of foreign currency.

Market clearing Letting $X_{i,t}^{\text{inv}}$ denote the stock of inventories at time t, $IM_{i,t}$ imports of the good that industry i produces, and $G_{i,t}$ sales to the government, market clearing for industry i requires that

$$X^{\text{inv}}_{i,t-1} + X_{i,t} + IM_{i,t} = \sum_j M_{i,j,t} + C_{i,t} + \sum_j I_{i,j,t} + G_{i,t} + \sum_d EX_{i,d,t} + X^{\text{inv}}_{i,t}. \tag{11}$$

Note that this framework is sufficiently general to precisely match industries' sales patterns in the data.

2.3 Estimating equations

We estimate the slope and curvature of the supply curve based on both the structural form and the reduced form.

2.3.1 Structural form

Linearizing the supply curve (10) around its $t-1$ values, and letting Δ denote the first difference operator, yields

$$\Delta \ln P^X_{i,t} = \mathcal{M}'\left(\ln u_{i,t-1}\right)\left(\Delta \ln X_{i,t} - \Delta \ln Q_{i,t}\right) + \Delta \ln mc_{i,t}. \tag{12}$$

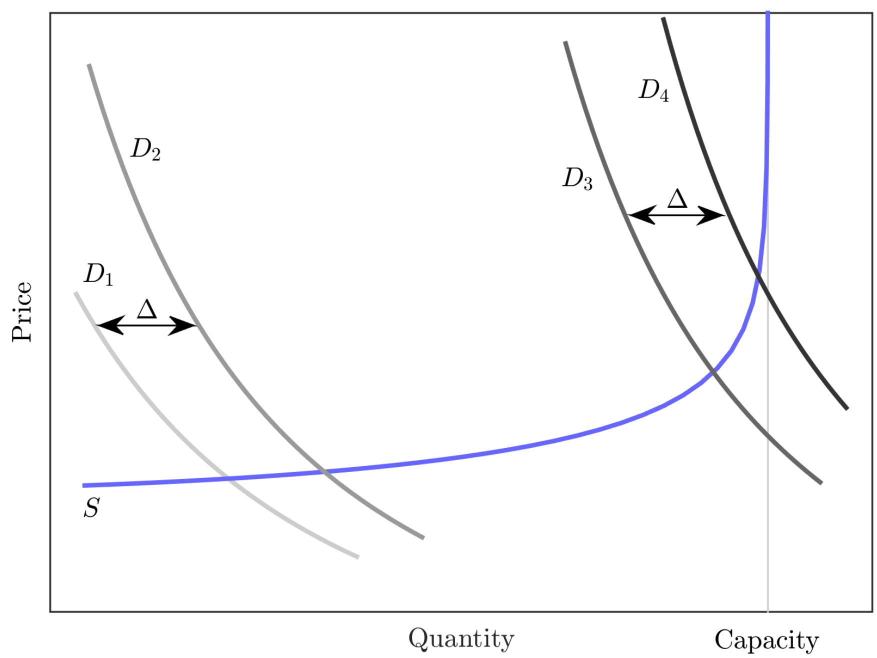
If the inverse supply elasticity $\mathcal{M}'$ is increasing in $\ln u_{i,t-1}$ the supply curve is convex. In that case, the initial equilibrium utilization rate determines how much prices and quantities respond locally to a demand shock. This intuition is illustrated in Figure (2).

2.3.2 Reduced form

We next define the following objects

$$\begin{aligned}
\Delta\xi_{i,t} &= \sum_j s^M_{i,j,t-1}\Delta \ln M_{j,t} + s^C_{i,t-1}\Delta \ln C_t + \sum_j s^I_{i,j,t-1}\Delta \ln I_{j,t} + s^G_{i,t-1}\Delta \ln G_{i,t} \qquad (13)\\
&\quad + \sum_d s^{\text{ex}}_{i,d,t-1}\Delta \ln EX_{d,t}\\
\Delta\pi_{i,t} &= \sum_j s^M_{i,j,t-1}\Delta \ln P^M_{j,t} + s^C_{i,t-1}\Delta \ln P^C_t + \sum_j s^I_{i,j,t-1}\Delta \ln P^I_{j,t} \qquad (14)\\
&\quad + \sum_d s^{\text{ex}}_{i,d,t-1}\Delta \ln P^{\text{ex},*}_{d,t}\\
\Delta e_{i,t} &= \sum_d s^{\text{ex}}_{i,d,t-1}\Delta \ln \mathcal{E}_{d,t}. \qquad (15)
\end{aligned}$$

Figure 2: The utilization rate as a sufficient statistic



In these expressions, $s_{i,j,t-1}$ denotes the share of industry i's overall sales to industry j, dated $t-1$. $\Delta\xi_{i,t}$ is an observable demand shifter which captures changes in industry i's customer's size. For instance, if industry j increases its demand for intermediates $M_{j,t}$ by one percent, industry i's demand rises, *ceteris paribus*, by $s^{M}_{i,j,t-1}$ percent. Similarly, a change in government purchases $\Delta \ln G_{j,t}$ affects industry i's demand.

$\Delta\pi_{i,t}$ reflects changes in demand due to changes in industry i's customers' prices. Continuing with the earlier example, industry i would, *ceteris paribus*, experience an increase in demand through substitution if downstream industry j's materials price index $P^{M}_{j,t}$ increased.

$\Delta e_{i,t}$ is the change of industry i's effective nominal exchange rate. $s^{\text{ex}}_{i,d,t-1}$ denotes the $t-1$ sales share of industry i to country d. Notice that $\Delta e_{i,t}$ varies by industry because existing trade linkages differentially expose industries to fluctuations of a common set of currencies. Further, this definition takes into account that some industries sell more of their goods abroad than others. A positive value of $\Delta e_{i,t}$ reflects a depreciation of the U.S. dollar relative to the relevant basket of foreign currencies. From the viewpoint of industry i which sets prices in dollars such a depreciation leads to an increase in demand through substitution towards the industry's product.

We next solve for the reduced form after log-linearizing around the equilibrium in $t-1$.

Proposition 2 (Reduced form)**.** *The industry's quantity, linearized around the equilibrium in $t-1$, is*

$$\begin{aligned}\Delta \ln X_{i,t} \quad = \quad & \beta_{\xi}\left(\ln u_{i,t-1}\right)\Delta\xi_{i,t}+\beta_{\pi}\left(\ln u_{i,t-1}\right)\Delta\pi_{i,t}+\beta_{e}\left(\ln u_{i,t-1}\right)\Delta e_{i,t} \\ & +\beta_{Q}\left(\ln u_{i,t-1}\right)\Delta\ln Q_{i,t}+\beta_{mc}\left(\ln u_{i,t-1}\right)\Delta\ln mc_{i,t} \\ & +\beta_{IM}\left(\ln u_{i,t-1}\right)\frac{\Delta IM_{i,t}}{X_{i,t-1}}+\beta_{inv}\left(\ln u_{i,t-1}\right)\frac{\Delta X_{i,t}^{inv}-\Delta X_{i,t-1}^{inv}}{X_{i,t-1}}+\omega_{i,t}^{X}.\end{aligned} \tag{16}$$

All coefficients are only functions of the log utilization rate $\ln u_{i,t-1}$ *and* $\beta_{\xi}>0$*,* $\beta_{\pi}>0$*,* $\beta_{e}>0$*,* $\beta_{mc}<0$*,* $\beta_{Q}>0$*,* $\beta_{IM}<0$*, and* $\beta_{inv}>0$*. The error term is a weighted average of changes in the unobserved demand shocks* $\omega_{i,t}^{C}$*,* $\omega_{i,j,t}^{I}$*,* $\omega_{i,j,t}^{M}$*, and* $\omega_{i,d,t}^{ex}$*.*

The equilibrium quantity is a function of all demand and supply shifters which—in this linearized version of the model—are $\Delta\xi_{i,t}$, $\Delta\pi_{i,t}$, $\Delta e_{i,t}$, $\Delta\ln Q_{i,t}$, $\Delta\ln mc_{i,t}$, $\Delta IM_{i,t}/X_{i,t-1}$, $\left(\Delta X_{i,t}^{\text{inv}}-\Delta X_{i,t-1}^{\text{inv}}\right)/X_{i,t-1}$ as well as $\omega_{i,t}^{X}$. The critical fact for our empirical analysis is that all coefficients β depend *only* on log utilization rates $\ln u_{i,t-1}$. This implies, for instance, that if the supply curve is convex, the elasticity $\beta_{\xi}\left(\ln u_{i,t-1}\right)$ is decreasing in the utilization rate so that the quantity response to a demand shock is larger if the initial utilization rate is low (see Figure 2). Detailed expressions for all coefficients in Proposition 2 are listed in Appendix B.

2.4 Discussion

Measurement of marginal costs In practice, the estimation of equations (12) and (16) is complicated by the fact that marginal costs are not observed. Further, subsuming marginal costs into the error term potentially leads to an omitted variable bias. An alternative is to proxy for marginal costs with unit variable costs which are observed. We briefly discuss the econometric implications of this strategy.

In our framework the industry's marginal costs differ from the industry's unit variable costs. This feature follows from the non-linear aggregation across varieties with aggregator (1).[5] Further, the wedge between unit variable cost and marginal cost is a function of utilization, that is, $\ln mc_{i,t}=\ln\frac{P_{i,t}^{v}X_{i,t}^{v}}{X_{i,t}}+\Omega\left(u_{i,t}\right)$. Substituting for marginal costs in equation (10) yields

$$\ln P_{i,t}^{X}=\mathcal{M}\left(\ln u_{i,t}\right)+\Omega\left(\ln u_{i,t}\right)+\ln\frac{P_{i,t}^{v}X_{i,t}^{v}}{X_{i,t}}. \tag{17}$$

[5] See Appendix B for details.

This expression makes clear that if, instead of marginal costs, unit variable costs are held constant, variation in u does not identify $\mathcal{M}'$, but $\mathcal{M}' + \Omega'$. A similar argument applies to $\mathcal{M}''$.

Proposition 3. *$\Omega \leq 0$, $\Omega' \leq 0$, and often $\Omega'' \leq 0$.*

Hence, empirical strategies that aim to estimate $\mathcal{M}'$ and $\mathcal{M}''$ based on equation (12), exhibit a downward bias (for both the slope and the curvature of $\mathcal{M}$) when marginal costs are proxied for with unit variable costs. It is thus possible that supply curves are upward-sloping and convex, but the researcher finds no evidence supporting this, even in large samples.

Capacity and anticipation effects cite Ramey 2011

Model misspecification Different sectors could have other determinants of markups. Partially address this by including changes in future prices.

Demand elasticities could be different across customers and time-varying.

2.5 Relationship to other frameworks

3 Empirical analysis

In this section we test empirically whether the data support the hypothesis that supply curves are convex at the industry level.

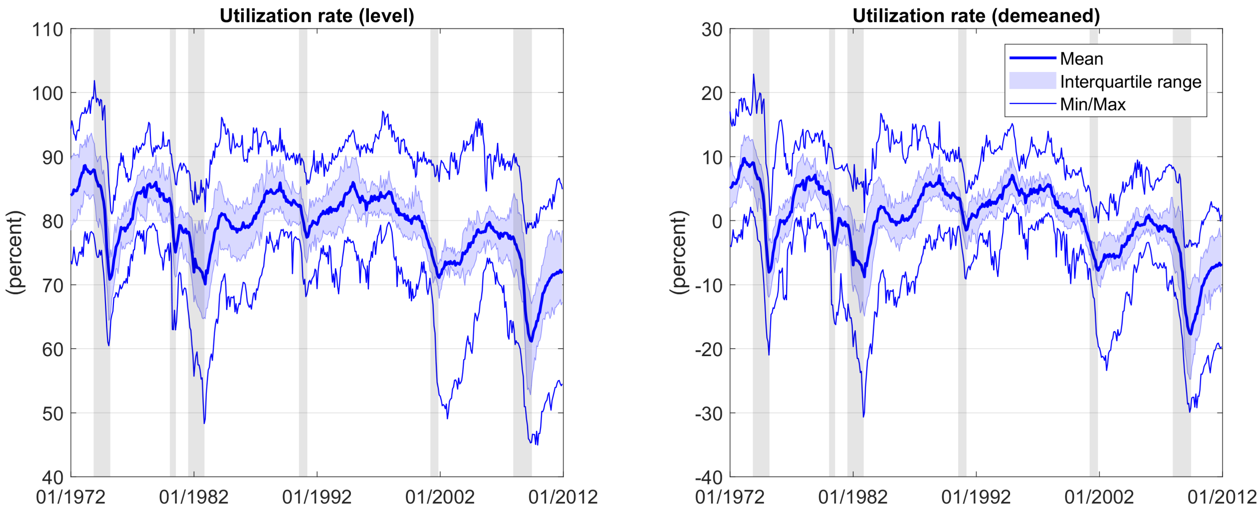
3.1 Data

3.1.1 Industrial production, capacity and utilization

We use the Federal Reserve Board's (FRB) measures of capacity and utilization. To obtain series for utilization, the FRB first constructs indexes of industrial production and capacity. The industrial production series are indexes of real gross output. The FRB's capacity indexes aim to capture the *sustainable maximum level of output*, that is, "the greatest level of output a plant [or industry] can maintain within the framework of a realistic work schedule after factoring in normal downtime and assuming sufficient availability of inputs to operate the capital in place".[6] The FRB's measure of capacity is primarily based on establishment-level

[6]See `https://www.federalreserve.gov/releases/g17/Meth/MethCap.htm`. A consequence of this definition is that utilization can exceed unity for short periods of time. In practice, this rarely happens. In our 3-digt NAICS sample from 1972 to 2011 only one industry (Primary Metal Manufacturing, NAICS 331) exceeded a utilization rate of 100 and only for two months (December of 1973 and January of 1974).

Figure 3: Capacity utilization at the industry level



Notes: The data are from the Federal reserve Board. The figure plots the mean, minimum, maximum and interquartile range of the capacity utilization series constructed using the FRB capacity utilization data and industrial production. The left panel shows the level of the utilization rate and the right panel industry-demeaned utilization series. Shaded areas represent NBER recessions.

utilization rates from the Survey of Plant Capacity (prior to 2007), the Quarterly Survey of Plant Capacity (from 2007 onwards), and measures of capital from the Annual Survey of Manufacturers.[7] As in our model, utilization is then calculated by dividing industrial production by capacity (see equation 9).

The left panel in Figure 3 illustrates the capacity utilization rates of the 21 3-digit NAICS manufacturing industries in our sample. As is clear from the figure, utilization rates display significant variation both in the cross-section and over time. Capacity utilization is strongly procyclical and experiences a mild downward trend towards the end of the sample, presumably reflecting the well-documented shrinkage of the U.S. manufacturing sector. An additional salient feature which we document in Appendix Table (D1) is that average utilization rates differ across industries. Since it is not clear to what extent these differences reflect measurement problems or industry-specific capacity choices we subtract from all utilization series their industry-specific means. The resulting series are shown in the right panel of Figure 3. We note that demeaning the utilization series does not drive our results.

[7]For further details on the data sources and methodology underlying of the capacity indexes, see Gilbert, Morin, and Raddock (2000) and `https://www.federalreserve.gov/releases/g17/About.htm`.

3.1.2 Additional data sources

We take data on prices, sales, input costs, and inventories from the NBER CES Manufacturing Industry Database. These data are constructed mainly from sources of the U.S. Census Bureau, the Bureau of Economic Analysis (BEA), and the Bureau of Labor Statistics (BLS) and provide a detailed picture of the U.S. manufacturing sector. For a description of this database, see Bartelsman and Gray (1996) and Becker, Gray, and Marvakov (2016). To obtain our measure of unit variable costs, $P^v_{i,t}X^v_{i,t}/X_{i,t}$ we sum production worker wages, costs of materials, and expenditures on energy and then divide by real gross output.

Our preferred measure of prices is a "deflator" constructed by dividing the nominal value of production (from the NBER CES) by the industrial production index (from the FRB). Relative to the price measure from the NBER CES database, this measure is consistent with the quantity measure (industrial production). We also show results using the price index from the NBER CES database. The results are very similar.

We calculate the sales shares $s_{i,j,t}$ from the Use Tables of the BEA's Input-Output accounts. For the sales shares to foreign countries, we complement these data with data on U.S. exports from the U.S. Census available from Peter Schott's website. The construction of $\Delta\xi_{i,t}$ and $\Delta\pi_{i,t}$ as given in equations (13) and (14) further requires data on quantity and price indexes. We use data from the following sources. First, for domestic sales of final goods we use data on personal consumption expenditures, equipment investment, and nonresidential fixed investment from the BEA's National Income and Product Accounts. Second, for intermediate sales to downstream industries, we use quantity and price indexes of industries' material use from the BEA's Industry Accounts. Third, for foreign quantity and price indexes we use real GDP and the GDP deflator in local currency from the United Nation's (UN) Statistics Division. The nominal exchange rate series for the calculation of $\Delta e_{i,t}$ (equation 15) also come from the UN's Statistics Division. To guarantee high data quality, we limit ourselves to countries that joined the Organisation for Economic Co-operation and Development (OECD) prior to year 2000 when constructing $\Delta\xi_{i,t}$, $\Delta\pi_{i,t}$, and $\Delta e_{i,t}$. Details on the data and the construction of these variables are available in Appendix C. Our sample is annual, includes all 21 3-digit NAICS manufacturing industries, and ranges from 1972 to 2011.

3.2 Empirical specifications

To estimate the structural form (12) and the reduced form (16), we approximate the coefficients linearly around the industry-specific mean $\ln \bar{u}_i$, so that, for instance,

$$\begin{aligned}\mathcal{M}'(\ln u_{i,t-1}) &= \mathcal{M}'(\ln \bar{u}_i) + \mathcal{M}''(\ln \bar{u}_i) \cdot (\ln u_{i,t-1} - \ln \bar{u}_i) \\ &\approx \mathcal{M}'(\ln \bar{u}_i) + \mathcal{M}''(\ln \bar{u}_i) \cdot (u_{i,t-1} - \bar{u}_i).\end{aligned}$$

Here, $u_{i,t-1} - \bar{u}_i$ are the demeaned utilization series plotted in the right panel of Figure 3. We always add a main effect of utilization $u_{i,t-1}$ to each specification in order to obtain the conventional interpretation of interaction terms.

We begin with estimating the structural form (12). In most specifications we use unit variable costs to proxy for marginal costs, although we keep in mind that doing so could generate downward biases in the slope and curvature of the supply curve (see Section 2.4). In a second step we estimate the reduced form for quantities (16).

3.2.1 Instruments and identification

Estimation of the slope and curvature of the supply curve requires an instrumental variable which shifts the demand curve and is excluded from the supply curve. When estimating the structural form (12), the instrument addresses the endogeneity of $\Delta \ln X_{i,t}$. Since we also estimate the reduced form, we use the same instrument for the supply shifter $\Delta\xi_{i,t}$ as defined in equation (13).[8]

We consider three different instruments in our empirical analysis. In all cases, the identification assumption requires that conditional on the control variables, the instrument is uncorrelated with the unobserved supply shifters. Whether this assumption is broadly satisfied depends on the instrument and the unobserved supply shocks (see Section 2.4). We emphasize that all instruments deliver comparable results.[9]

World import demand The first instrument aims to use variation in foreign demand to estimate the supply curve. Letting $Y_{d,t}$ denote GDP in destination d, our World Import

[8] We recognize the ambiguity of the term reduced from in the context of this paper. We will use the term to refer to the equilibrium quantity and price as a function of all shifters.

[9] We have also considered a fourth instrument based on defense spending. Since the first stage of this instrument is uniformly weak and the estimates noisy, we do not report the results.

Demand instrument is

$$\text{inst}_{i,t}^{\text{WID}} = \sum_d s_{i,d,t-1}^{\text{ex}} \Delta \ln Y_{d,t}. \tag{18}$$

To better understand the identification assumption, we decompose the change in foreign GDP into a common and a destination-specific component, $\Delta \ln Y_{d,t} = \Delta \ln Y_t^{\text{com}} + \Delta \ln Y_{d,t}^{\text{spec}}$. Letting $\bar{s}_{d,t-1}^{\text{ex}}$ denote the average export share of all manufacturing industries to destination d, the variation of the World Import Demand instrument can be split into three components,

$$\text{inst}_{i,t}^{\text{WID}} = \Delta \ln Y_t^{\text{com}} \sum_d s_{i,d,t-1}^{\text{ex}} + \sum_d \bar{s}_{d,t-1}^{\text{ex}} \Delta \ln Y_{d,t}^{\text{spec}} + \sum_d \left(s_{i,d,t-1}^{\text{ex}} - \bar{s}_{d,t-1}^{\text{ex}} \right) \Delta \ln Y_{d,t}^{\text{spec}}. \tag{19}$$

The first term on the right hand side captures variation common to all foreign countries. Since this variation reflects the "global" business cycle and may affect the industries in the sample through channels not captured by other controls, we control for it by interacting a time fixed effect with the foreign sales share of industry i, $\sum_d s_{i,d,t-1}^{\text{ex}}$.

The second term on the right hand side weighs destination-specific changes in GDP with the average export share. Since most of our specifications will include a time fixed effect—in addition to the time fixed effect interacted with the foreign sales share—this variation will be purged as well. Hence, the identifying variation of this instrument entirely comes from the third term, $\sum_d \left(s_{i,d,t-1}^{\text{ex}} - \bar{s}_{d,t-1}^{\text{ex}} \right) \Delta \ln Y_{d,t}^{\text{spec}}$, reflecting destination-specific changes in GDP which are weighted with the deviations of sales shares from the average. Instruments of this type have been used by Hummels et al. (2014), among others.

Shea's instrument Our second instrument builds on Shea (1993a,b), who argues that material demand of industry j from industry i is more likely to constitute an "exogenous" demand shock if industry i does not rely on inputs from industry j. Recall that $s_{i,j,t-1}^{M}$ denotes industry i's sales share to industry j. Further, let $s_{i,j,t-1}^{\text{cost}}$ denote industry i's materials cost share from industry j. Our version of the Shea-instrument is then

$$\text{inst}_{i,t}^{\text{Shea}} = \sum_j s_{i,j,t-1}^{M} \mathbb{1}\left\{ s_{i,j,\tau}^{M} > 10 \cdot s_{i,j,\tau}^{\text{cost}} \; \forall \tau \right\} \Delta \ln M_{j,t}. \tag{20}$$

The condition in the indicator function requires that in every period for which we have data the sales share from industry i to j exceeds ten times industry i's materials cost share from j. It aims at generating a high degree of relevance while containing the degree of endogeneity (see Shea (1993a,b) for details).

Note that this instrument only partially addresses concerns on reverse causality. Supply shocks that affect industry i and spill over to industry j through input linkages, and shocks which are correlated across industries must be dealt with by adding appropriate control variables. We will do so by controlling for unit variable costs to address the former concern and by adding time fixed effects and time fixed effects interacted with the foreign sales share to address the latter.

The exchange rate We further use a purified change in an industry's effective exchange rate (equation 15) to identify the slope and curvature of the supply curve. A dollar depreciation relative to the relevant basket of foreign currencies makes U.S.-produced goods cheaper for foreign customers. If firms in the U.S. set prices in U.S. dollars (as 97 percent of U.S. exporters do, see Gopinath and Rigobon (2008)), such depreciations materialize as outward shifts in demand. A one percent depreciation of the effective exchange rate raises demand by the value of the demand elasticity (σ in Section 2.2).

Analogously to the World Import Demand instrument, we purge changes in the effective exchange rate $\Delta e_{i,t}$ by decomposing the nominal exchange rate into a common and destination-specific component $\Delta \ln \mathcal{E}_{d,t} = \Delta \ln \mathcal{E}_t^{com} + \Delta \ln \mathcal{E}_{d,t}^{spec}$. This decomposition can be implemented by regressing the observed changes in exchange rates on a set of time fixed effects. In our sample, the R^2 of this regression is 28.3 percent, implying that 28.3 percent of changes in the dollar value of foreign currencies are common to all foreign currencies.

Similar to the World Import Demand instrument in equation (19), we decompose the effective exchange rate into three components,

$$\Delta e_{i,t} = \Delta \ln \mathcal{E}_t^{\text{com}} \sum_d s_{i,d,t-1}^{\text{ex}} + \sum_d \bar{s}_{d,t-1}^{\text{ex}} \Delta \ln \mathcal{E}_{d,t}^{\text{spec}} + \sum_d \left(s_{i,d,t-1}^{\text{ex}} - \bar{s}_{d,t-1}^{\text{ex}} \right) \Delta \ln \mathcal{E}_{d,t}^{\text{spec}}. \quad (21)$$

When the specification includes a time fixed effect and a time fixed effect interacted with the foreign sales share, the identifying variation of the exchange rate instrument is limited to destination-specific exchange rate changes which are weighted with the deviations of sales shares from the average (the third term on the right hand side).

As Amiti, Itskhoki, and Konings (2014) emphasize, most exporters also import and hence dollar depreciations raise the cost of intermediates inputs. To prevent this channel from confounding our interpretation of dollar depreciations as demand shocks, we control for unit variable costs as suggested by the model in most specifications.

3.2.2 Additional notes on identification

To be completed.

3.3 Results

We begin with presenting the estimates of the structural form and subsequently turn to the estimates of the reduced form.

3.3.1 Estimates of the structural form

Table 1 shows the estimates of the supply curve when we impose linearity. Specification (1) begins with an estimate of the inverse supply elasticity without controls. As expected, the slope estimate is insignificantly different from zero since unobserved supply shocks confound the estimation. When we add the change in unit variable costs in specification (2), the R-squared rises to 87 percent and the estimate of the slope coefficient becomes positive and significant. Adding unit variable costs to the specification partially addresses the simultaneity problem arising from the existence of unobserved supply shocks because they explain such a large fraction of the variation. Specification (3) further adds the change in capacity to the equation. As expected from equation (12), the coefficient is negative. Further, the slope coefficient rises to 0.13. In specification (4) we also add industry fixed effects, time fixed effects and time fixed effects interacted with the industry's export share. The estimate of the inverse supply elasticity rises to 0.17. When we simultaneously use the World Import Demand instrument, Shea's instrument, and the effective exchange rate instrument in specification (5), we obtain a slope estimate of 0.21. This estimate is greater than the OLS estimate and precisely estimated. The first stage F-statistic is 12.97. We pass Hansen's J test with a p-value of 0.418.

We next relax the assumption of linearity and allow the inverse supply elasticity to depend on last period's utilization rate (see Figure 2). Table 2 specification (1) shows the OLS estimates. The main effect is 0.17 and the interaction term is negative and insignificant. We next use the World Import Demand instrument, Shea's instrument, and the effective exchange rate to instrument for the main effect (as in specification (5) of Table 1), and additionally use the World Import Demand instrument interacted with the demeaned utilization rate $u_{i,t-1} - \bar{u}_i$ for the interaction term. As specification (2) shows, the interaction term becomes positive and significant. This suggests that the inverse supply elasticity is increasing in the initial capacity utilization rate.

Table 1: Estimates of the linear model

Dependent variable: $\Delta \ln P_{i,t}$

Estimator	OLS	OLS	OLS	OLS	2SLS
Instruments	-	-	-	-	WID, Shea, $\Delta e_{i,t}$
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln X_{i,t}$	-0.09	0.08	0.13	0.17	0.21
	(0.08)	(0.02)	(0.02)	(0.02)	(0.06)
$\Delta \ln Q_{i,t}$			-0.16	-0.12	-0.14
			(0.03)	(0.04)	(0.05)
$\Delta \ln P^v_{i,t} X^v_{i,t} / X_{i,t}$		0.91	0.90	0.89	0.89
		(0.02)	(0.02)	(0.03)	(0.03)
R-squared	0.010	0.873	0.880	0.910	0.909
First stage F					12.97
Hansen J (p-value)					0.418
Fixed Effects	no	no	no	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

Little changes when we alternatively instrument for the interaction term with Shea's instrument interacted with $u_{i,t-1} - \bar{u}_i$ as shown in specification (3). The estimate of the inverse supply elasticity also changes little when we use the effective exchange rate instead (specification 4), although this instrument is potentially weak. When we use all three instruments interacted with the demeaned utilization rate, the coefficient on the interaction term is 0.89 and precisely estimated (specification 5).

Because the exchange rate instrument is potentially weak, we drop it from our set of instruments both of the main effect and the interaction term. This leaves us with the WID instrument and Shea's instrument. As specification (6) demonstrates, both the main effect and the interaction term are positive and highly significant ($p < 0.01$). Under the identification assumption that the WID and Shea's instrument are orthogonal to unobserved supply shocks, supply curves at the industry level are increasing and convex. We next turn to a number of robustness checks.

Table (3) specification (1) reports the estimates when we drop unit variable costs from the set of controls. As discussed in Section (2.4), including this control potentially leads to a downward bias in both the slope and the curvature of the supply curve. On the other hand,

Table 2: Estimates of the non-linear model

Dependent variable: $\Delta \ln P_{i,t}$

Estimator	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Instrument(s):						
Main effect		WID, Shea, $\Delta e_{i,t}$				WID, Shea
Interaction ($\cdot\, (u_{i,t-1} - \bar{u}_i)$)		WID	Shea	$\Delta e_{i,t}$	all	WID, Shea
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln X_{i,t}$	0.17	0.23	0.23	0.23	0.23	0.22
	(0.02)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)
$\Delta \ln X_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-0.33	0.88	0.73	0.83	0.89	0.90
	(0.24)	(0.31)	(0.37)	(0.54)	(0.29)	(0.29)
$u_{i,t-1} - \bar{u}_i$	0.01	0.02	0.02	0.02	0.02	0.02
	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
$\Delta \ln Q_{i,t}$	-0.11	-0.19	-0.19	-0.18	-0.18	-0.18
	(0.04)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-0.32	-1.05	-0.96	-1.01	-1.05	-1.05
	(0.43)	(0.40)	(0.40)	(0.33)	(0.40)	(0.40)
$\Delta \ln P^v_{i,t} X^v_{i,t} / X_{i,t}$	0.89	0.89	0.89	0.89	0.89	0.89
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
R-squared	0.910	0.904	0.905	0.905	0.904	0.904
SW First stage F: main effect		17.62	14.82	24.07	8.99	14.61
SW First stage F: interaction		14.83	16.12	5.31	22.49	18.23
Hansen J (p-value)					0.748	0.576
Fixed Effects	yes	yes	yes	yes	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

this control explains approximately 87 percent of the variation in price changes and is thus critical for obtaining precise results. Dropping unit variable costs from the controls indeed increases the estimate of the slope and curvature. Due to the increase in error variance, however, the interaction term becomes insignificant. To obtain precise estimates we continue the empirical analysis including unit variable costs as a control, but keep in mind that this control potentially leads to downward biases.

The estimating equation (12) suggests that the slope coefficient is the negative of the coefficient on capacity (and similarly for the interaction term). Inspection of the estimates in Table 2 suggests that this cross-coefficient restriction broadly holds. However, this observation also raises the concern that the estimates of the slope and curvature of the supply curve are dependent on the inclusion of capacity and its interaction with the utilization rate. Specifications (2) and (3) drop these controls from the equation. The estimates of the slope and curvature fall slightly but remain positive and significant.

We next add a number of additional controls to the equation. Specification (4) adds the percent change of the industry's price from t to $t+1$. Extensions of the model with sticky prices suggest that this variable should capture the firm's expectations of changes in future marginal costs. Adding this control has essentially no effect on the estimates. In specification (5) we include an interaction term of changes in unit variable costs with the utilization rate. This variable is positive and highly significant, suggesting that pass-through of cost shocks into prices may be stronger when capacity utilization is high. The slope and curvature of the supply curve change little.

A number of studies have used the capacity utilization rate to forecast inflation—with varying conclusions (see Shapiro (1989) and Stock and Watson (1999)). In line with Shapiro's (1989) conclusion, the coefficient on the utilization rate is small and insignificant in all specifications of Table 2. Yet, when we add the squared utilization rate in specification (6) of Table (3), both the linear and the squared term are positive and significant. Consistent with the convex shape of the supply curve, the utilization rate appears to have predictive power for price changes when utilization is high, but not when it is low. The slope and curvature estimates of the supply curve increase slightly when this additional control is included.

In specification (7) we add $\Delta\pi_{i,t}$ as defined in equation (14) to address the possibility that strategic complementarities lead the industry to change prices when its competitors change prices. While the control is highly significant, the slope and curvature estimates remain broadly unchanged. In our preferred specification (8) we simultaneously include all additional controls. The slope estimate is 0.30 and the coefficient on the interaction term 1.31. The estimates are similar when we replace the left hand side variable (the change in the implicit "deflator") with changes in the price index (specification 9).

Table 3: Robustness of the non-linear model

Dependent variable: $\Delta \ln P_{i,t}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta \ln X_{i,t}$	0.54	0.21	0.19	0.22	0.26	0.28	0.23	0.30	0.27
	(0.15)	(0.06)	(0.05)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)
$\Delta \ln X_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	1.09	0.75	0.64	0.89	0.88	1.34	1.01	1.31	1.07
	(0.82)	(0.27)	(0.25)	(0.30)	(0.30)	(0.34)	(0.29)	(0.38)	(0.30)
$u_{i,t-1} - \bar{u}_i$	0.39	0.01	-0.02	0.02	0.00	0.07	0.02	0.05	0.00
	(0.09)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)
$\Delta \ln Q_{i,t}$	-0.64	-0.16		-0.18	-0.20	-0.24	-0.19	-0.25	-0.16
	(0.19)	(0.08)		(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-1.73			-1.04	-0.96	-2.47	-1.25	-2.21	-1.69
	(0.86)			(0.40)	(0.38)	(0.56)	(0.40)	(0.70)	(0.55)
$\Delta \ln P^v_{i,t} X^v_{i,t} / X_{i,t}$		0.89	0.89	0.89	0.90	0.90	0.85	0.87	0.91
		(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.09)
$\Delta \ln P_{i,t+1}$				0.00				0.01	-0.10
				(0.02)				(0.02)	(0.06)
$\Delta \ln P^v_{i,t} X^v_{i,t} / X_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$					1.12			0.74	1.67
					(0.30)			(0.34)	(0.48)
$(u_{i,t-1} - \bar{u}_i)^2$						1.08		0.80	0.54
						(0.24)		(0.26)	(0.24)
$\Delta \pi_{i,t}$							0.27	0.25	-0.09
							(0.09)	(0.09)	(0.14)
R-squared	0.516	0.905	0.904	0.904	0.908	0.903	0.905	0.906	0.858
SW First stage F: main effect	14.08	15.21	14.39	13.58	13.69	11.14	14.72	10.97	6.97
SW First stage F: interaction	18.09	24.59	30.40	17.18	15.59	12.30	20.12	11.57	9.30
Hansen J (p-value)	0.440	0.504	0.522	0.580	0.533	0.501	0.492	0.432	0.291
Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

Figure 4: Non-parametric estimates

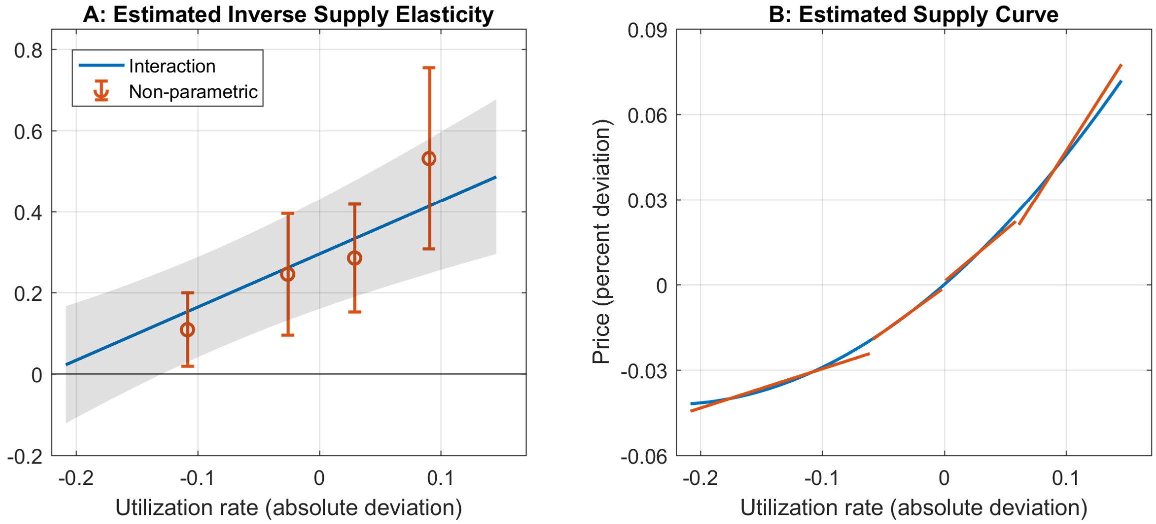


Figure 4 Panel A shows non-parametric estimates of the inverse supply elasticity. We allow the slope coefficient to differ depending on whether the utilization rate of the previous period ($u_{i,t-1} - \bar{u}_i$) was below -0.06 (approximately the 15th percentile), between -0.06 and 0, between 0 and 0.06, and above 0.06 (approximately the 85th percentile). The non-parametric estimates broadly align with those based on the interaction term.

Based on specification (8) of Table (3) the figure also demonstrates that below the fifth percentile of the utilization distribution ($u_{i,t-1} - \bar{u}_i = -0.13$), the estimated inverse supply elasticity is below 0.12 and statistically indistinguishable from zero. This contrasts to a value of 0.42 at the 95th percentile ($u_{i,t-1} - \bar{u}_i = 0.10$). The non-parametric estimate for the highest utilization rates suggests an even large value of 0.53, but it is imprecisely estimated. Panel B of Figure 4 plots the quadratic and the partially-linear approximation of the supply curve.

3.3.2 Estimates of the reduced form

We next turn to the estimates of the reduced form (equation 16). The object of primary interest in this section is the coefficient on $\Delta\xi_{i,t}$, which captures how much production responds to a generic demand shock. This coefficient, for instance, can be interpreted as the industry-level fiscal multiplier under the assumption that the model is correctly specified. That is, it captures by how much the industry's production expands when the government purchases goods worth one dollar. We are also interested in the coefficient on $\Delta e_{i,t}$, which measures the production response to a depreciation of the U.S. dollar.

Specifications (1) and (2) of Table 4 show these estimates of a linear model without and with controlling for changes imports and inventory accumulation. All coefficients have the expected sign, and those on $\Delta\xi_{i,t}$ and $\Delta e_{i,t}$ are highly significant. Since imports and inventory accumulation are presumably correlated with the error, it is not clear as to whether to include them in the regression. Since doing so has little effect on the other coefficients, we proceed with including them.

In specification (3) we add industry fixed effects, time fixed effects, and time fixed effects interacted with the foreign sales share. While the coefficients on $\Delta\xi_{i,t}$ and $\Delta e_{i,t}$ change little, the standard error on the coefficient of $\Delta e_{i,t}$ more than doubles. The reason is that, taken together, these fixed effects explain approximately 94 percent of the variation in the exchange rate variable (the time fixed effects interacted with the foreign sales share alone explain 92.6 percent). An implication of this is that the main effect of the effective exchange rate will be imprecisely estimated in all specifications.

When we instrument for $\Delta\xi_{i,t}$ with the WID instrument, Shea's instrument, or both (specifications 4 to 6), the estimates remain very stable between 0.8 and 0.9. This finding is not obvious because shocks that affect one industry could spill over to other industries via input-output linkages and ultimately feed back to the first industry, thereby creating a simultaneity problem. Shea's instrument ameliorates this issue by considering changes in downstream demand without such feedback. Yet, the 2SLS estimates are very similar to the OLS estimates, suggesting that, conditional on the controls, such feedback loops do not drive much of the variation in industry's production.

Specification (1) in Table 5 presents OLS estimates of the non-linear reduced form. This specification is based on equation (16) and all coefficients are allowed to linearly depend on the utilization rate. To preserve space, we only report the coefficients on the variables of interest. The coefficient on the interaction term $\Delta\xi_{i,t}\cdot(u_{i,t-1}-\bar{u}_i)$ is negative and significant. Similarly, the interaction term associated with the exchange rate $\Delta e_{i,t}\cdot(u_{i,t-1}-\bar{u}_i)$ is negative and significant. Consistent with the implications of a convex supply curve, demand shocks appear to stimulate production more when the utilization rate is initially low.

The interaction term $\Delta\xi_{i,t}\cdot(u_{i,t-1}-\bar{u}_i)$ becomes more negative when we estimate the specification by 2SLS. With the WID instrument (specification 2), the coefficient becomes -4.11, and with Shea's instrument it becomes -2.58 (specification 3). When we combine both instrument in specification (4), the coefficient is -2.98. Hence, the initial utilization rate determines how much production responds to demand shocks.

Table 4: Estimates of the linear reduced form

Dependent variable: $\Delta \ln X_{i,t}$

Estimator	OLS	OLS	OLS	2SLS	2SLS	2SLS
Instrument(s)				WID	Shea	WID, Shea
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\xi_{i,t}$	0.89	0.91	0.84	0.90	0.82	0.82
	(0.08)	(0.09)	(0.11)	(0.41)	(0.14)	(0.14)
$\Delta\pi_{i,t}$	-0.02	-0.01	-0.05	-0.06	-0.05	-0.05
	(0.11)	(0.12)	(0.12)	(0.12)	(0.11)	(0.11)
$\Delta e_{i,t}$	1.89	2.03	1.66	1.57	1.70	1.69
	(0.59)	(0.44)	(1.02)	(1.29)	(0.98)	(0.99)
$\Delta \ln Q_{i,t}$	0.68	0.78	0.64	0.63	0.65	0.65
	(0.07)	(0.07)	(0.08)	(0.11)	(0.08)	(0.08)
$\Delta \ln P^v_{i,t} X^v_{i,t}/X_{i,t}$	-0.08	-0.05	-0.12	-0.12	-0.12	-0.12
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
$\left(\Delta X^{\text{inv}}_{i,t} - \Delta X^{\text{inv}}_{i,t-1}\right)/X_{i,t-1}$		0.10	0.03	0.03	0.03	0.03
		(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
$\Delta IM_{i,t}/X_{i,t-1}$		-0.08	0.14	0.11	0.15	0.15
		(0.08)	(0.08)	(0.18)	(0.09)	(0.09)
R-squared	0.689	0.703	0.810	0.810	0.810	0.810
First stage F				10.06	543.39	315.73
Hansen J (p-value)						0.843
Fixed Effects	no	no	yes	yes	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

We next consider a number of robustness checks. In specifications (1) of Table 6, we drop unit variable costs and its interaction with the utilization rate from the regression. The estimates change very little. The estimates are also robust to alternatively dropping the change in capacity and its interaction from the regression (specification 2). In specifications (3) and (4) we add the change in future prices and a lagged dependent variable. Both of these variables are significant, but including them in the regression barely affects the estimates. We further estimate a specification with all second order terms. Specification (5) includes squares and interactions of all control variables. The interaction term $\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ falls slightly in absolute magnitude (to -2.27), but it remains significant at the one percent level. Finally, specification (6) presents estimates when we include all second order terms in addition to the change in future prices and a lagged dependent variable. In this preferred specification

Table 5: Estimates of the non-linear reduced form

Dependent variable: $\Delta \ln X_{i,t}$

Estimator	OLS	2SLS	2SLS	2SLS
Instrument(s):				
Main effect		WID, Shea		
Interaction ($\cdot (u_{i,t-1} - \bar{u}_i)$)		WID	Shea	WID, Shea
	(1)	(2)	(3)	(4)
$\Delta\xi_{i,t}$	0.81	0.74	0.74	0.74
	(0.09)	(0.12)	(0.13)	(0.13)
$\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-1.72	-4.11	-2.58	-2.98
	(0.64)	(0.91)	(1.02)	(0.74)
$\Delta e_{i,t}$	1.16	1.12	1.22	1.19
	(0.96)	(0.90)	(0.90)	(0.90)
$\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-21.42	-20.77	-21.00	-20.95
	(4.72)	(4.98)	(4.62)	(4.67)
$(u_{i,t-1} - \bar{u}_i)$	-0.28	-0.24	-0.27	-0.26
	(0.06)	(0.04)	(0.05)	(0.05)
R-squared	0.834	0.828	0.833	0.836
SW First stage F: main effect		372.51	355.04	277.05
SW First stage F: interaction		18.50	35.39	44.70
Hansen J (p-value)		0.903	0.886	0.623
Other controls	yes	yes	yes	yes
Fixed Effects	yes	yes	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

coefficient on $\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ is -2.14 and that on $\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ is -16.85. Hence, that the production response depends on the initial utilization rate is robust to including or dropping a large number controls.

Figure 5 plots the estimated inverse supply elasticity. The estimates are based on specification (6) of Table 6. The figure also plots non-parametric estimates that allow for different production responses depending on whether the initial utilization rate is below -0.06 , between -0.06 and 0, between 0 and 0.06, and above 0.06. These estimates align well with those based on the interaction term. As Panel A demonstrates, production responds by approximately twice as much when the initial utilization rate is below the fifth percentile (-0.13) than when it is above the 95th percentile (0.10). The exchange rate response drops to zero at high utilization rates (Panel B).

Table 6: Robustness of the reduced form

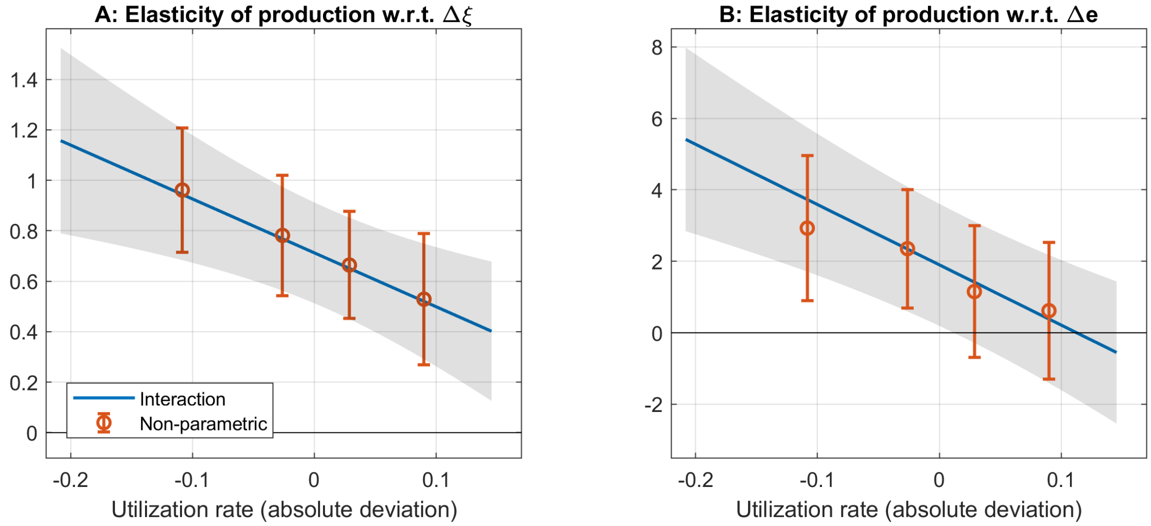
Dependent variable: $\Delta \ln X_{i,t}$

Estimator	2SLS					
Instruments						
Main effect	WID, Shea					
Interaction ($\cdot (u_{i,t-1} - \bar{u}_i)$)	WID, Shea					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\xi_{i,t}$	0.73	0.81	0.73	0.74	0.73	0.71
	(0.13)	(0.14)	(0.13)	(0.13)	(0.12)	(0.12)
$\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-2.92	-2.83	-2.98	-2.79	-2.27	-2.14
	(0.73)	(0.56)	(0.78)	(0.75)	(0.82)	(0.86)
$\Delta e_{i,t}$	1.15	1.55	1.08	1.41	1.94	1.90
	(0.86)	(1.23)	(0.90)	(0.88)	(1.01)	(1.03)
$\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-20.68	-22.52	-20.47	-19.45	-18.82	-16.85
	(4.72)	(5.80)	(4.54)	(4.27)	(5.51)	(5.08)
$(u_{i,t-1} - \bar{u}_i)$	-0.27	-0.07	-0.26	-0.27	-0.22	-0.23
	(0.05)	(0.06)	(0.05)	(0.05)	(0.07)	(0.07)
$\Delta \ln P_{i,t+1}$			0.07			0.07
			(0.02)			(0.02)
$\Delta \ln X_{i,t-1}$				0.12		0.15
				(0.04)		(0.05)
R-squared	0.831	0.787	0.833	0.835	0.845	0.850
SW First stage F: main effect	274.35	282.38	275.91	235.86	250.44	218.58
SW First stage F: interaction	52.66	47.15	43.59	46.40	46.11	46.45
Hansen J (p-value)	0.757	0.580	0.596	0.623	0.892	0.716
Other controls	yes	yes	yes	yes	yes	yes
Drop $\Delta \ln P^v_{i,t} X^v_{i,t}/X_{i,t}$ and interaction	yes	no	no	no	no	no
Drop $\Delta \ln Q_{i,t}$ and interaction	no	yes	no	no	no	no
Second order terms	no	no	no	no	yes	yes
Fixed Effects	yes	yes	yes	yes	yes	yes

Notes: The estimates are based on equation (12). Driscoll-Kraay standard errors are reported in parentheses.

We finally ask the question whether the estimates of the structural form and the reduced form are consistent with one another. How much production responds to an outward shift in demand depends both on the supply elasticity and the elasticity of demand. In Figure 6 we plot this elasticity of production with respect to the demand shock $\Delta\xi_{i,t}$. The figure shows both the direct estimate based on the reduced form and the response implied by

Figure 5: Non-parametric estimates of the reduced form



the estimated supply elasticity. We plot this latter response for three alternative demand elasticities, $\sigma = 1$, 2, and 3. As is clear from the figure, both estimates are broadly consistent with on another. The fact that both estimates are consistent with one another suggests that it is indeed the slope of the supply curve that determines locally how much industries respond to demand shocks—and not an alternative mechanism such as rationing.

4 Applications

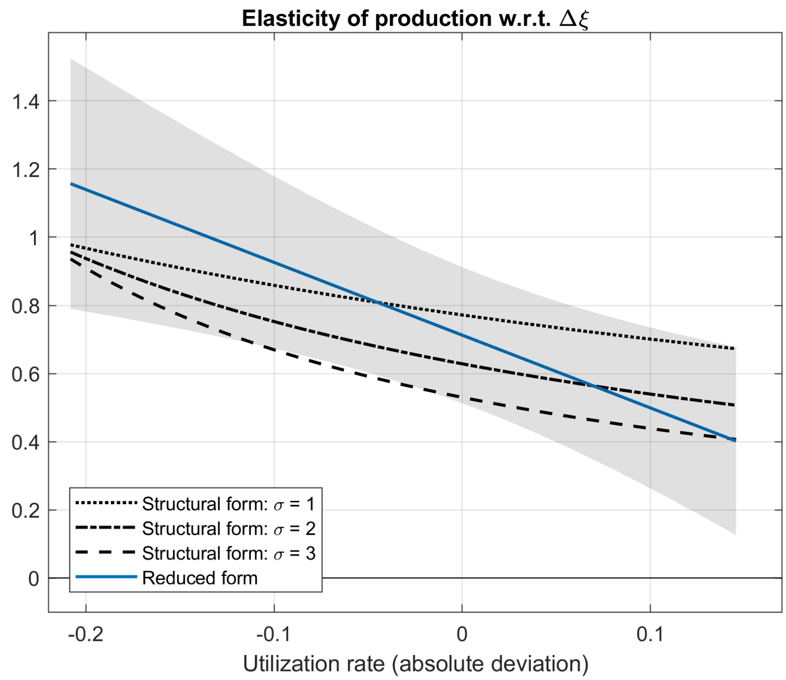
4.1 State-Dependent Multipliers

To be completed.

4.2 A non-linear Phillips Curve

To be completed.

Figure 6: The production response to demand shocks



5 Conclusion

This paper studies whether supply curves are convex in the short-run. To guide our empirical analysis we develop a putty-clay model in which short-run capacity constraints can generate a convex supply curve at the industry level. Using a sufficient statistics approach, we estimate the model and find strong support for the convexity of supply curves. Industries with low initial capacity utilization rates expand production much more after dollar depreciations or defense spending shocks than industries that produce close to their capacity limit. Further, prices rise after such demand shocks only if the initial level of capacity utilization is high. Our evidence of convex supply curves at the industry level has a number of applications, including that that capacity constraints are a likely candidate for generating state-dependent responses to shocks, etc.

To be completed.

References

Alvarez-Lois, Pedro P. 2004. "Capacity constraints, idiosyncratic demand uncertainty and the dynamics of inflation." *Economics Letters* 83 (1):15 – 21.

———. 2006. "Endogenous capacity utilization and macroeconomic persistence." *Journal of Monetary Economics* 53 (8):2213 – 2237.

Amiti, Mary, Oleg Itskhoki, and Jozef Konings. 2014. "Importers, Exporters, and Exchange Rate Disconnect." *American Economic Review* 104 (7):1942–1978.

Auerbach, Alan J. and Yuriy Gorodnichenko. 2012. "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4 (2):1–27.

———. 2013a. "Fiscal Multipliers in Recession and Expansion." In *Fiscal Policy after the Financial Crisis*, NBER Chapters. National Bureau of Economic Research, Inc, 63–98.

———. 2013b. "Output Spillovers from Fiscal Policy." *American Economic Review* 103 (3):141–146.

Baqaee, David Rezza and Emmanuel Farhi. 2017. "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem." Working Paper 23145, National Bureau of Economic Research.

Bartelsman, Eric J. and Wayne Gray. 1996. "The NBER Manufacturing Productivity Database." Working Paper 205, National Bureau of Economic Research.

Baum, Anja, Marcos Poplawski-Ribeiro, and Anke Weber. 2012. "Fiscal Multipliers and the State of the Economy." Imf working papers, International Monetary Fund.

Becker, Randy, Wayne Gray, and Jordan Marvakov. 2016. "NBER-CES Manufacturing Industry Database: Technical Notes." Tech. rep.

Bils, Mark and Jang-Ok Cho. 1994. "Cyclical factor utilization." *Journal of Monetary Economics* 33 (2):319–354.

Burnside, Craig and Martin Eichenbaum. 1996. "Factor-Hoarding and the Propagation of Business-Cycle Shocks." *American Economic Review* 86 (5):1154–1174.

Cooley, Thomas F, Gary D Hansen, and Edward C Prescott. 1995. "Equilibrium Business Cycles with Idle Resources and Variable Capacity Utilization." *Economic Theory* 6 (1):35–49.

Fagnart, Jean-Francois, Omar Licandro, and Henri R. Sneessens. 1997. "Capacity utilization and market power." *Journal of Economic Dynamics and Control* 22 (1):123–140.

Fagnart, Jean-Franois, Omar Licandro, and Franck Portier. 1999. "Firm Heterogeneity, Capacity Utilization and the Business Cycle." *Review of Economic Dynamics* 2 (2):433–455.

Fazzari, Steven M., James Morley, and Irina Panovska. 2015. "State-dependent effects of fiscal policy." *Studies in Nonlinear Dynamics & Econometrics* 19 (3):285–315.

Gilbert, Charles, Norman J. Morin, and Richard D. Raddock. 2000. "Industrial production and capacity utilization: recent developments and the 1999 annual revision." *Federal Reserve Bulletin* (Mar):20–33.

Gilchrist, Simon and John C. Williams. 2000. "Putty-Clay and Investment: A Business Cycle Analysis." *Journal of Political Economy* 108 (5):928–960.

Gopinath, Gita and Roberto Rigobon. 2008. "Sticky Borders." *Quarterly Journal of Economics* 123 (2):531–575.

Gorodnichenko, Yuriy and Matthew Shapiro. 2011. "Using the Survey of Plant Capacity to Measure Capital Utilization." Working papers, Center for Economic Studies, U.S. Census Bureau.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman. 1988. "Investment, Capacity Utilization, and the Real Business Cycle." *American Economic Review* 78 (3):402–417.

Hansen, Gary D. and Edward C. Prescott. 2005. "Capacity constraints, asymmetries, and the business cycle." *Review of Economic Dynamics* 8 (4):850–865.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang. 2014. "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data." *American Economic Review* 104 (6):1597–1629.

Huo, Zhen, Andrei A. Levchenko, and Nitya Pandalai-Nayar. 2019. "The Global Business Cycle: Measurement and Implications." Tech. rep.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2017. "Large and State-Dependent Effects of Quasi-Random Monetary Experiments." Nber working papers, National Bureau of Economic Research, Inc.

Kaplan, Greg and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica* 82 (4):1199–1239.

Kuhn, Florian and Chacko George. 2017. "Business Cycle Implications of Capacity Constraints under Demand Shocks." Working Paper, .

Lo, Ming Chien and Jeremy Piger. 2005. "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model." *Journal of Money, Credit and Banking* 37 (5):865–886.

Michaillat, Pascal. 2014. "A Theory of Countercyclical Government Multiplier." *American Economic Journal: Macroeconomics* 6 (1):190–217.

Owyang, Michael T., Valerie A. Ramey, and Sarah Zubairy. 2013. "Are Government Spending Multipliers Greater during Periods of Slack? Evidence from Twentieth-Century Historical Data." *American Economic Review* 103 (3):129–34.

Parker, Jonathan A. 2011. “On Measuring the Effects of Fiscal Policy in Recessions.” *Journal of Economic Literature* 49 (3):703–718.

Peersman, Gert and Frank Smets. 2005. “The Industry Effects of Monetary Policy in the Euro Area.” *Economic Journal* 115 (503):319–342.

Ramey, Valerie A. and Sarah Zubairy. 2018. “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data.” *Journal of Political Economy* 126 (2):850–901.

Shapiro, Matthew D. 1989. “Assessing the Federal Reserve’s Measures of Capacity Utilization.” *Brookings Paper on Economic Activity* (1):181–241.

Shea, John. 1993a. “Do Supply Curves Slope Up?” *The Quarterly Journal of Economics* 108 (1):1–32.

———. 1993b. “The Input-Output Approach to Instrument Selection.” *Journal of Business & Economic Statistics* 11 (2):145–155.

Stock, James H. and Mark W. Watson. 1999. “Forecasting inflation.” *Journal of Monetary Economics* 44 (2):293–335.

Weise, Charles L. 1999. “The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach.” *Journal of Money, Credit and Banking* 31 (1):85–108.

A Appendix: Capacity utilization at the plant level

Add figure on qualitative responses.

We begin with presenting several basic facts on plants' capacity utilization using microdata from the Quarterly Survey of Plant Capacity Utilization (QSPC). The survey is conducted by the U.S. Census Bureau and funded jointly by the Federal Reserve Board and the Department of Defense. The sample is drawn from all U.S. manufacturing and publishing plants with 5 or more production employees. Among other things, establishments are asked about the market value of their *actual production* and the estimated market value of their *full production capacity*. Respondents are asked to construct this estimate under the following assumptions: 1) only the current functional machinery and equipment is available, 2) normal downtime, 3) labor, materials, and other non-capital inputs are fully available, 4) a realistic and sustainable shift and work schedule, and 5) that the establishment produces the same product mix as its current production.[10] Capacity utilization rates are then obtained by dividing the market value of actual production by the estimate of full capacity production.

Figure 1 plots the density estimates of utilization rates for the years 2007, 2009 and 2011. Utilization rates display substantial cross-sectional variation and three facts emerge from the figure. First, a significant fraction of establishments produces at full capacity. In all three years, a discrete mass of establishments bunches at a utilization rate of unity. Second, an even larger share of plants produces below their reported capacity, at times far below: Utilization rates between 0.2 and 0.5 are not uncommon. Finally, capacity utilization rates and the fraction of firms with utilization rates of unity are highly procyclical. In 2007 a large fraction of plants produced at full capacity and the density displays a mode at around 0.8. By 2009 the distribution has shifted to the left and has a modal point at approximately 0.5. The 2011 density reflects partial recovery relative to 2009 but utilization rates are still below those of 2007.

We next turn to the question why establishments produce at levels below their capacity. Respondents of the QSPC are asked: "If this plant's actual production in the current quarter was less than full production capacity, mark (X) the primary reasons." Possible answers include "Insufficient supply of materials", "Insufficient orders", "Insufficient supply of local labor force/skills", and others. Multiple answers are permitted. It turns out that the vast majority of plants produce below capacity because they are not able to sell their products. For the time period from 2013q1 to 2017q2 for which public data is available, 79.7 percent of plant managers cite insufficient orders as the main reason for producing below capacity. The second most cited option is chosen by 10.0 percent of respondents (insufficient supply of local labor force/skills).

[10] A survey form of the Quarterly Survey of Plant Capacity Utilization is available at `https://bhs.econ.census.gov/bhs/pcu/watermark_form.pdf`.

B Appendix: Model Extensions

C Data Appendix

C.1 Sample and data sources

Our baseline sample is annual and includes all 21 3-digit NAICS manufacturing industries. It ranges from 1972 to 2011.

The sales shares to foreign countries $s_{i,d,t}$ are constructed based on sales to all countries that joined the OECD prior to year 2000. These are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, the Republic of Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States.

Export data The data on exports are from the U.S. Census and are available from Peter Schott's website `http://faculty.som.yale.edu/peterschott/sub_international.htm`. This data is available with SIC industry codes between 1972 and 1997, and with NAICS industry codes thereafter. We use the NBER CES SIC4 to NAICS6 concordance based on sales weights to convert the SIC codes into NAICS equivalents and then aggregate to the 3-digit NAICS level.

D Appendix: Additional Results

Table D1: Summary Statistics of Utilization Rates by 3-digit NAICS Manufacturing Industries

Industry	NAICS	p10	Median	p90	Mean	S.D.	Skewness	Kurtosis	Durable
Food	311	79.6	82.3	85.2	82.4	2.4	0.3	2.5	no
Beverage and Tobacco Products	312	68.3	79.2	83.0	77.3	5.3	-0.5	2.1	no
Textile Mills	313	68.3	82.0	89.5	79.8	8.6	-0.8	3.2	no
Textile Product Mills	314	69.8	82.3	90.4	80.9	8.3	-0.8	3.2	no
Apparel	315	71.0	80.2	84.2	79.0	4.9	-0.9	3.4	no
Leather and Allied Products	316	59.3	74.9	82.1	72.8	8.8	-1.2	3.7	no
Wood Products	321	63.8	79.2	85.2	77.1	8.4	-1.2	4.6	yes
Paper	322	81.4	87.6	91.4	86.9	4.2	-0.2	2.4	no
Printing and Related Support Activities	323	72.2	82.7	89.3	81.3	7.6	-1.0	3.8	no
Petroleum and Coal Products	324	77.3	87.1	92.6	85.7	5.8	-0.7	2.8	no
Chemicals	325	72.1	77.8	83.1	77.7	4.3	-0.4	2.3	no
Plastics and Rubber Products	326	71.4	83.7	89.6	82.4	7.2	-0.9	3.3	no
Nonmetallic Mineral Products	327	62.3	77.2	84.0	75.3	9.2	-1.6	5.3	yes
Primary Metals	331	68.2	79.6	89.5	79.3	9.3	-0.7	3.4	yes
Fabricated Metal Products	332	71.7	77.7	84.4	77.4	5.7	-0.2	3.1	yes
Machinery	333	67.6	78.9	87.0	77.8	7.8	-0.2	2.5	yes
Computer and Electronic Product	334	70.1	79.0	84.2	78.2	5.7	-1.0	4.0	yes
Electrical Equipment, Appliances, and Components	335	73.2	82.8	90.6	82.6	6.7	-0.2	2.6	yes
Transportation Equipment	336	66.4	75.6	81.5	74.4	6.1	-1.0	4.1	yes
Furniture and Related Products	337	68.0	77.8	84.1	76.8	7.4	-0.2	3.9	yes
Miscellaneous	339	72.8	76.9	79.7	76.3	3.1	-0.5	3.1	yes
All		70.0	79.8	88.6	79.1	7.6	-0.8	4.4	

Source: Federal Reserve Board