Convex Supply Curves*

Christoph E. Boehm[1] and Nitya Pandalai-Nayar[2]

[1]University of Texas at Austin
[2]University of Texas at Austin and NBER

May 2020

Abstract

We provide evidence that industries' supply curves are convex. To guide our empirical analysis, we develop a model, in which capacity constraints at the plant level generate convex supply curves at the industry level. The industry's capacity utilization rate is a sufficient statistic for the supply elasticity. Using data on capacity utilization and three different instruments, we estimate the supply curve and find robust evidence for an economically sizable degree of convexity. The nonlinearity we identify has several macroeconomic implications: Responses to shocks are state-dependent, the Phillips curve is convex, and higher welfare costs of business cycles than Lucas (1987).

JEL Codes: E22, E32, E52, E62, F44

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, David Baqaee, Oli Coibion, Javier Cravino, Aaron Flaaen, Oleg Itskhoki, Andrei Levchenko, Valerie Ramey, Chris Sims, and Jaume Ventura as well as with seminar participants at UT Austin, Princeton, Wisconsin, the Federal Reserve Board, the St Louis Fed, Tsinghua, Texas A&M, SED Annual Meetings, ASSA Annual Meetings, the West Coast International/Macro Workshop and Hong Kong University. 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. More recent research has emphasized the role of nonlinearities, but the empirical evidence on their importance is often limited. When the evidence exists, it is frequently controversial. For instance, there is a lack of consensus as to whether the fiscal multiplier varies with the state of the business cycle.[1]

In this paper we document robust evidence for one particular nonlinearity—that industries' supply curves are convex. Our econometric estimates suggest that the curvature of industries' supply curves is economically large. Industries that produce close to capacity respond to a positive demand shock by raising their prices by six times as much as industries that produce far below their capacity. At the same time, their production response is cut in half. We demonstrate in a multi-sector general equilibrium model that convex supply curves have sizable and policy relevant implications in the aggregate, including a countercyclical fiscal multiplier, a convex Phillips curve, and greater welfare costs of business cycles than standard models.

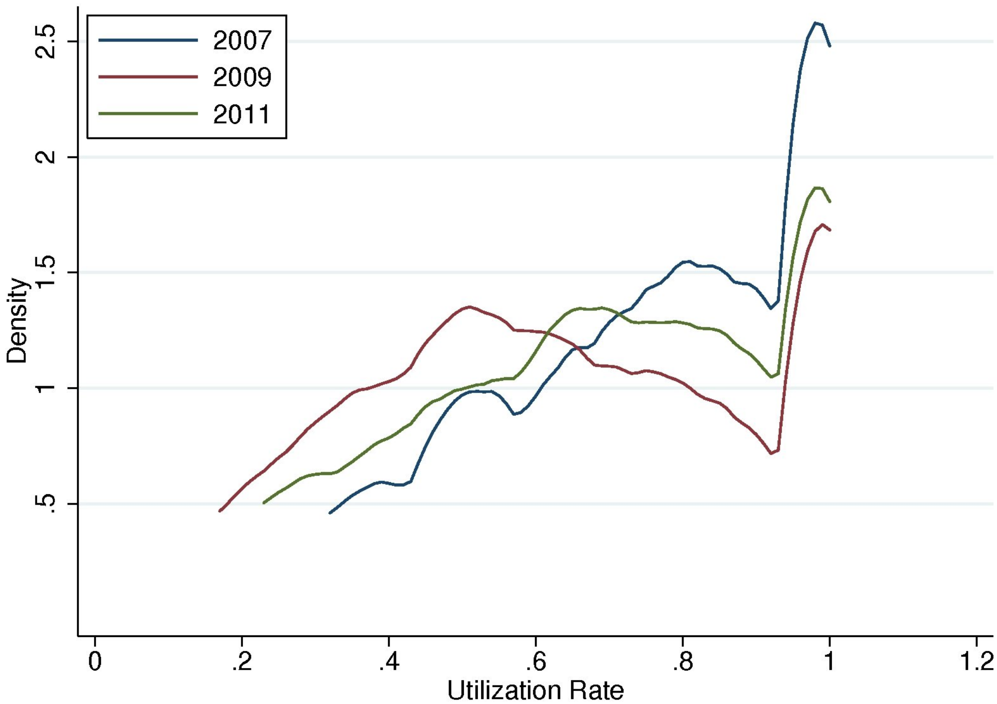
We begin by deriving our estimating equation in a model in which capacity constraints at the plant level generate convex supply curves at the industry level. Micro data from the Quarterly Survey of Plant Capacity (QSPC) suggests that capacity constraints indeed limit plants' production. As Figure 1 illustrates, a significant fraction of U.S. manufacturing establishments produces at "full capacity" as defined in the QSPC or, equivalently, at a capacity utilization rate of one. These plants presumably have limited room for increasing production in the short run and are likely to raise their price when confronted with an increase in demand.

We formalize this logic by assuming that firms invest into a set of factors that are fixed in the short run and, once chosen, determine the firm's maximum production capacity. When the demand for a firm's good materializes sufficiently high, production becomes constrained by capacity. Constrained firms' production is locally unresponsive to shocks because any changes in demand are absorbed in the markup. The framework permits simple aggregation to the industry level, where it generates a supply curve that is convex *in logarithms*.[2] Its key implication is that an industry's capacity utilization rate is a sufficient statistic for its supply elasticity. The model builds on prior work by Fagnart, Licandro, and Portier (1999), is qualitatively consistent with plant-level facts on capacity utilization, and delivers the estimating equation for our empirical analysis.

Our main contribution is to provide empirical evidence that industry's supply curves are convex. The intuition of our empirical strategy is as follows. As Figure 1 shows, the fraction of capacity constrained plants varies over time. In 2007 many plants produced at full capacity. Subsequently,

[1]There are exceptions for which nonlinearities have been firmly established. These include, for instance, the work on lumpy price adjustment and investment behavior. For evidence on the state-dependence of the fiscal multiplier, see Auerbach and Gorodnichenko (2012; 2013) who argue that the multiplier is likely large during downturns, and Ramey and Zubairy (2018) who find that such state dependence is small or nonexistent.

[2]The model's supply curve is convex *in logarithms*, implying that the inverse supply elasticity is increasing in output. This contrasts with commonly used production functions such as constant elasticity of substitution (CES). When parameterized with decreasing returns to scale, these production functions generate cost functions that are linearly increasing in logs and thus have a constant (inverse) supply elasticity. Throughout this paper, we refer to supply curves as convex when they are convex in logs.



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. We describe the data and discuss additional results in Appendix A.

Figure 1: Densities of plant capacity utilization

the Great Recession induced a leftward shift of the distribution of utilization rates and fewer plants were constrained by capacity. By 2011, the distribution of utilization rates had partially recovered. If constrained plants respond differently to demand shocks than unconstrained plants, the fraction of constrained plants will affect an industry's pricing and production responses to such shocks. Our estimation exploits this variation across time and industries by allowing industries' price responses to depend on their initial capacity utilization rate. We show formally in our model that in a regression of prices on quantities an interaction term with the utilization rate identifies the curvature of the supply curve.

We estimate the supply curve using three alternative instruments to trace out its slope and curvature. First, we use a version of the World Import Demand instrument (Hummels et al., 2014). This instrument assumes that appropriately purified changes in foreign imports 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 large downstream industries. Third, we consider changes in industries' effective exchange rates. Conditional on holding industries' costs constant, depreciations in the exchange rate stimulate demand from abroad. All three instruments deliver comparable results.

The estimates suggest that supply curves are highly elastic at low levels of capacity utilization. At low capacity utilization rates, 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 between 0.41 and 0.61 at high levels of capacity utilization. We also directly estimate production responses to demand shifts (the reduced form). We find that production responds twice as much to the same sized demand shock for industry-year observations below the 15th percentile when compared to observations above the 85th percentile. These estimates are based on an annual sample from 1972 to 2011, which covers all 3-digit U.S. manufacturing industries. We use the measures of capacity utilization from the Federal Reserve Board (FRB), which are close empirical analogues to the corresponding object in the model.[3]

To quantify the aggregate implications of convex supply curves in general equilibrium, we embed the estimated industry equilibrium framework into a nonlinear network model of 71 industries. In our benchmark calibration the fiscal multiplier is countercyclical and varies between 0.55 and 0.68 over the business cycle. This degree of state-dependence is not trivial, but smaller than estimates by Auerbach and Gorodnichenko (2012, 2013). The partial elasticity of real marginal costs with respect to output—which plays a critical role for determining the slope of the Phillips curve in models with price rigidities—is procyclical and drops by approximately 20 percent during the Great Recession. This finding may partially explain the "missing disinflation" observed at the time and is relevant for stabilization policy that relies on generating inflation (e.g. Christiano, Eichenbaum, and Rebelo, 2011). Finally, the welfare costs of business cycles are approximately twice as high as in Lucas (1987).

Related literature Early work on capacity goes back to Chamberlin (1933) and Robinson (1933). Cassels (1937) noted that definitions of capacity require an assumption on which factors of production are held fixed. In line with this argument, today's measures of capacity from the QSPC and its predecessor, the Survey of Plant Capacity (SPC), assume that a plant's capital stock is fixed, while all other factors can be freely varied. In the 1950s and 1960s, Smithies (1957), Klein (1960), and others began to introduce capacity and utilization into business cycle theories, and studied theoretical questions on measurement of capacity. A recurring topic highlighted by this literature is that engineering-based concepts of capacity are unsatisfactory from the point of view of economic decision making, since they disregard firms' economic incentives. The SPC and the QSPC explicitly address this issue by defining capacity as the technically *and economically* sustainable level of production.[4]

Standard neoclassical production theory does not feature a generally accepted concept of capacity. In the absence of clear theoretical guidance, early series of capacity utilization were often *ad-hoc* (Perry, 1973), and were later discontinued or the underlying methodology changed. Shapiro (1989) provided a negative assessment of the Federal Reserve Board's (FRB) measures of capacity and utilization at the time, since the evidence appeared to be inconsistent with the intuition that "[c]apacity is best thought of as the level of output where the marginal cost curve becomes steep" (Shapiro, 1989, p. 204). Our evidence is consistent with this intuition, and thus provides an *ex-post* validation of the FRB's current measures. We note, however, that it is neither theoretically nor empirically clear whether the steepness of the supply curve close to capacity arises

[3]Our estimates qualify as "identified moments" in the sense of Nakamura and Steinsson (2018); they are portable to other environments, including to models without a concept of capacity.

[4]For instance, the QSPC asks respondents about their "full production capability" under the assumption that "the number of shifts, hours of operation and overtime pay that can be sustained under normal conditions and a realistic work schedule in the long run."

from increases in marginal costs or markups. Following Shapiro's critique, the FRB revised its methodology of measuring capacity in 1990. The revised series are consistently available from 1972 onwards and are primarily based on the SPC and the QSPC. Corrado and Mattey (1997), Gilbert, Morin, and Raddock (2000), and Morin and Stevens (2004) discuss the methodology in detail, we provide a brief summary in Section 3. Of course, the FRB's primary objective is to use its measures of capacity utilization to gauge inflationary pressures. Among many others, Stock and Watson (1999) find capacity utilization to be a useful predictor of inflation.

The concept of capacity utilization is distinct from—although 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. These include Greenwood, Hercowitz, and Huffman, 1988, Bils and Cho, 1994, Cooley, Hansen, and Prescott, 1995, Burnside and Eichenbaum, 1996, Gilchrist and Williams, 2000, Hansen and Prescott, 2005, among others. Our framework builds on Fagnart, Licandro, and Portier (1999), who explicitly model capacity utilization and study the business cycle implications in a linearized model economy. This paper instead studies the *nonlinear* implications of capacity constraints. Our findings also demonstrate that of the model's concept of capacity is operationalizable; the model's definition of capacity is a close theoretical analogue of the FRB's current measures and empirical evidence supports the prediction of a convex supply curve. Overall, the nonlinear implications of capacity utilization and related notions of slack have received relatively little attention. Two important 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 a single sector general equilibrium model.

Lastly, our paper is related to prior work that estimates production functions and the degree of returns to scale. Some early evidence suggested increasing returns to scale (e.g. Hall, 1990). Subsequent work by Burnside, Eichenbaum, and Rebelo (1995), Basu and Fernald (1997), Bils and Klenow (1998), and others, rejected increasing returns in favor of constant or decreasing returns. Perhaps the most closely related paper is Shea (1993a), who found that supply curves in many industries slope up. His benchmark estimate of the inverse supply elasticity on a pooled sample is 0.18. When we disregard second order terms and estimate a linear specification, we obtain a very similar value of 0.23. Estimation techniques of production functions in industrial organization typically assume that the production function is Cobb-Douglas (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg, Caves, and Frazer, 2015). This assumption rules out that production functions are convex in logs.

Roadmap 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 empirical results in Section 3. We discuss the aggregate general equilibrium implications of our findings in Section 4. Section 5 concludes.

2 Theoretical framework

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

For maximum clarity, the model in this section assumes that prices are flexible. In Appendix C we derive our estimating equations in a version of the model with sticky prices.

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 within an industry. 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 production decision. If demand materializes sufficiently high, production will be constrained by capacity. The supply curve's convexity is the result of capacity-constrained plants charging higher markups. To preserve the model's tractability, we assume constant marginal costs at the plant level, although we note that increasing marginal costs could also generate a convex supply curve at the industry level.

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,

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

$\theta \in (1, \infty)$ 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 that these shocks are drawn independently and identically from distribution G with $\mathbb{E}[\omega] = 1$ and $\mathbb{E}[\omega^2] < \infty$.

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

$$y_t(j) = \omega_t(j) Y_t \left(\frac{p_t^y(j)}{P_t^Y} \right)^{-\theta} \tag{2}$$

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

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

2.1.2 Intermediate goods producers

Consistent with Figure 1, we assume that a firm's capacity can limit production in the short run. Each 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

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

That is, production y_t is linear in 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 $y_t = z_t F\left(\kappa_t, 1\right) v_t$ 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, discount future profits at rate r, and maximize the present value of profits. We allow firm's investment to be subject to possibly non-convex adjustment costs $\phi\left(x, k\right)$. The firm's Bellman equation is then

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

where the maximization is subject to

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

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

as well as equations (2), (4), and (5). Equation (7) is the capacity constraint and (8) is the standard capital accumulation equation. We assume that productivity z only has an industry-specific and an aggregate (i.e. economy wide), but no firm-specific component. This assumption limits the degree of heterogeneity in the model and allows us to analytically aggregate output and prices to the industry level.[5]

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.

[5] In this version of the model without idiosyncratic productivity shocks, all firms choose identical values for the capital stock k and the maximum of variable inputs $\bar{v}$. While introducing idiosyncratic productivity shocks makes the model intractable analytically, we conjecture that our key conclusion that the supply curve is convex in logs remains unchanged. As we discuss below, the two key assumptions that lead to the convexity of the supply curve are that 1) varieties are imperfectly substitutable and 2) the maximum output of each variety is predetermined within the period.

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,

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

where ρ is the multiplier on the capacity constraint (7). In this baseline version of the model rising markups are the key mechanism generating a convex supply curve at the industry level.

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

$$Y\left(q_t, \bar{\omega}_t\right) = q_t \left(\left(\bar{\omega}_t\right)^{-\frac{\theta-1}{\theta}} \int_0^{\bar{\omega}_t} \omega dG\left(\omega\right) + \int_{\bar{\omega}_t}^{\infty} \left(\omega\right)^{\frac{1}{\theta}} dG\left(\omega\right) \right)^{\frac{\theta}{\theta-1}}. \tag{9}$$

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

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

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

Further, an industry's capacity utilization rate is defined as the ratio of actual production to full capacity production,

$$u_t := \frac{Y\left(q_t, \bar{\omega}_t\right)}{Q\left(q_t\right)}. \tag{10}$$

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 (10) has the following properties:*

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

Proof. See Appendix B. □

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. For the empirical analysis, this property is crucial because u_t is observable, while the threshold variety $\bar{\omega}$ is not.[6]

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

$$\ln P_t^Y = \mathcal{M}(\ln u_t) + \ln(mc_t). \tag{11}$$

This (inverse) supply curve depends on the industry's marginal costs mc_t, and the industry's *log average markup* $\mathcal{M}$. This markup is only a function of the industry's utilization rate. Since variation in u_t is observable, this equation is the starting point for our empirical analysis. Note that 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:*

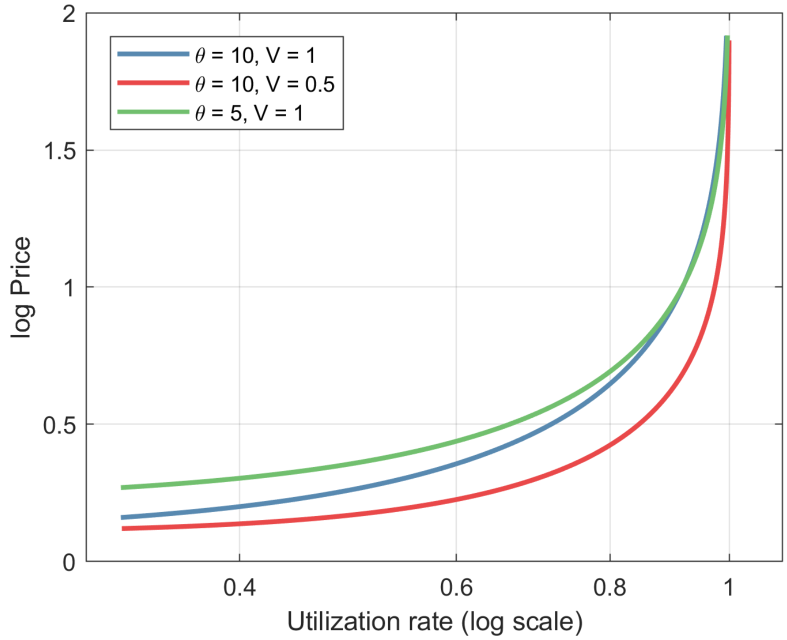
1. $\mathcal{M}' \geq 0$
2. $\lim_{u\to 0} \mathcal{M}(\ln u) = \ln \frac{\theta}{\theta-1}$, $\lim_{u\to 1} \mathcal{M}(\ln u) = \infty$
3. $\lim_{u\to 0} \mathcal{M}'(\ln u) = 0$, $\lim_{u\to 1} \mathcal{M}'(\ln u) = \infty$
4. *Without further restrictions on G, the sign of $\mathcal{M}''$ is generally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

Proof. See Appendix B. □

Because $\mathcal{M}$ is increasing in utilization everywhere, the industry's supply curve (11) is upward-sloping. In contrast to standard models, the industry's markup rises when production Y_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. Hence, the model predicts under all parameterizations that the supply curve is flat for low and steep for high utilization rates.

At a more fundamental level, the shape of the supply curve follows from the assumptions that 1) the maximum output for each variety is predetermined within the period (the putty-clay assumption), and 2) varieties are imperfectly substitutable ($\theta < \infty$). For the aggregating firm, these two assumptions imply that raising output necessitates substitution towards less productive varieties, whose production is not capacity constrained. This raises the industry's price index.

[6] Lemma 1 also implies that the utilization rate as defined in equation (10) should correlate highly with the fraction of capacity-constrained plants. This prediction can be verified in the data. As Figure 1 illustrates, the fraction of constrained plants strongly co-moves with the average utilization rate. Measuring utilization rates directly as a weighted fraction of constrained plants is undesirable for a number of reasons. First, it is plausible that firms begin raising their markup when approaching capacity and not only when hitting the constraint. Second, and as discussed in Morin and Stevens (2004), the uncorrected survey data exhibit a "cyclical bias", which the Federal Reserve removes when constructing their estimates. Third, doing so would break the relationship between the utilization rate, output, and capacity (equation (10)). Since our census project does not have data for all years of our sample in Section 3, it is also practically not feasible.



Notes: The figure provides illustrative examples of supply curve (11). The parameterization is chosen as follows: G is log-normal with unit mean and variance V, marginal costs mc are set to 1.

Figure 2: Supply curves

For suppliers of these varieties, imperfect substitutability allows them to charge markups—which increase with demand if the capacity constraint binds. If all producers are constrained, their markups grow and the industry's price becomes infinite.

For most economically meaningful choices of G, the log markup $\mathcal{M}(\ln u)$ is convex everywhere. Figure 2 shows parameterized examples of supply curve (11) when G is log-normal. A lower demand elasticity θ and a lower variance of the demand shock increase the curvature of the supply curve. The intuition is as follows. To raise industry output, the aggregating firm can only increase purchases of inputs from suppliers whose production is not constrained by capacity. For lower values of θ this substitution is costlier, thus raising the curvature of the supply curve. Further, for this log-normal example, a lower variance implies a higher density of varieties around the mean. Near the mean, the aggregating firm must therefore *to a greater degree* substitute away from varieties that become constrained at the margin. This drives up the industry's price index P_t^Y (see proof of Proposition 1 in Appendix B for details).

As we demonstrate in Figure B1 of Appendix B, it is possible to construct examples in which $\mathcal{M}(\ln u)$ is locally concave. We note, however, that these examples rely on placing little to no density at the center of the distribution of demand shocks G (a bimodal distribution), which makes these cases less likely to be of economic interest. Further, whether $\mathcal{M}$ is convex in the relevant range of utilization is ultimately an empirical question, which we address below.

2.1.5 Estimating equation

Letting Δ denote the first difference operator and adding industry subscripts i, linearization of the supply curve (11) around its $t-1$ values yields

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

We then approximate $\mathcal{M}'\left(\ln u_{i,t-1}\right)$ linearly around the industry-specific mean $\ln \bar{u}_i$ so that

$$\mathcal{M}'\left(\ln u_{i,t-1}\right) = \mathcal{M}'\left(\ln \bar{u}_i\right) + \frac{\mathcal{M}''\left(\ln \bar{u}_i\right)}{\bar{u}_i} \cdot \left(u_{i,t-1} - \bar{u}_i\right).$$

Plugging this expression back into equation (12), and adding a constant and an error term, gives our estimating equation for the structural form,

$$\begin{aligned} \Delta \ln P_{i,t}^{Y} = {} & \alpha + \beta_Y \Delta \ln Y_{i,t} + \beta_{Yu} \Delta \ln Y_{i,t} \cdot \left(u_{i,t-1} - u_i\right) + \beta_u \left(u_{i,t-1} - u_i\right) \\ & + \beta_Q \Delta \ln Q_{i,t} + \beta_{Qu} \Delta \ln Q_{i,t} \cdot \left(u_{i,t-1} - u_i\right) + \beta_{mc} \Delta \ln mc_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{13}$$

The primary object of interest is the coefficient on the interaction term $\Delta \ln Y_{i,t} \cdot \left(u_{i,t-1} - \bar{u}_i\right)$. If this coefficient is positive, the supply curve is convex. To obtain the conventional interpretation of the interaction terms, we always include a main effect of the demeaned utilization rate $u_{i,t-1} - \bar{u}_i$. The error term should be thought of as resulting from model misspecification and/or measurement error in the data.

Relative to the ad-hoc estimation of supply curves, our model provides two key insights. First, in this framework, the initial utilization rate is a sufficient statistic for the supply elasticity. This feature allows us to estimate the curvature of the supply curve with an interaction term $\Delta \ln Y_{i,t} \cdot \left(u_{i,t-1} - \bar{u}_i\right)$, rather than a squared term in output $\left(\Delta \ln Y_{i,t}\right)^2$. As we will demonstrate below, using the interaction term has a critical econometric advantage. While we can find strong instruments for the interaction term, our instruments are weak for the squared term.

Second, unlike conventional specifications of supply curves, equation (13) includes the change in capacity as an endogenous supply shifter. All else equal, greater production capacity implies that fewer firms in the industry are constrained. This reduces the industry's price index and increases the industry's output (see Proposition 2 below). Further, it is likely that firms occasionally adjust capacity in response to new information they receive, including to expected changes in future demand. Since such changes in capacity shift the supply curve, they generate downward-biased slope and curvature estimates when subsumed into the error term. Instruments that shift the demand curve are unlikely to address this bias. If an instrument identifies persistent changes in demand, the industry will adjust its capacity in response, and the supply curve will shift. In the next section, we will therefore include the change in capacity as a control variable. We will also verify that the slope and curvature estimates fall when we drop the change in capacity from the set of controls.

2.1.6 Discussion

Why do plants produce below capacity? In the model, plants produce below capacity in equilibrium because they cannot sell more of their product at the desired price. This feature is consistent with the data from the QSPC. For the time period from 2013q1 to 2018q2 for which public QSPC data is available, 78.4 percent of plant managers cite "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 this sense, the model is qualitatively consistent with the micro data. Details on these qualitative survey responses are available in Appendix A.

Adjustment costs and the evolution of $Q_{i,t}$ The supply curve (11) was derived without taking a stance on the form of adjustment costs ϕ, which affect firms' capacity choices and ultimately the evolution of the industry's capacity $Q_{i,t}$. This fact implies that, conditional on observing $Q_{i,t}$, the industry's supply curve is independent of the precise form of adjustment costs that firms face. Since $Q_{i,t}$ is measured in the data and imposing less structure increases generality, we view this independence as a desirable property.

Anticipation effects A related point is that, for fixed $\Delta \ln mc_{i,t}$ and $\Delta \ln Y_{i,t}$, news about the future affects the industry's price changes $\Delta \ln P^Y_{i,t}$ *only* through its effect on $\Delta \ln Q_{i,t}$—which is explicitly forward-looking. This implies that, conditional on observing $\Delta \ln Q_{i,t}$, firms' private information, which is not observable to the econometrician, should not pose a problem for identification. As discussed earlier, the leading example here is an anticipated change in future demand. Since the change in future demand affects current prices only through its effect on capacity, the control $\Delta \ln Q_{i,t}$ captures all relevant information.

Markups versus marginal costs In this simple model, the positive slope and curvature of supply curve (11) is entirely driven by increasing markups, and the marginal cost term mc_t is independent of the utilization rate (see equation (6)). In reality, this may not be the case. As firms add additional shifts, for instance, overtime premia may raise the marginal cost of production (see, e.g. Bresnahan and Ramey, 1994). When taking the model to the data, we therefore prefer to interpret the dependence of $\mathcal{M}$ on the utilization rate as potentially reflecting both changes in firms' markups as well as changes in their marginal costs. Since marginal costs are not observed, we cannot separately identify markups from marginal costs. As we discuss below, we will control for unit variable costs (a proxy for marginal costs), but these will not capture a possible dependence of marginal costs on the utilization rate.

For the empirical analysis, the precise mechanism generating the slope and curvature of the supply curve is immaterial. As long as we correctly isolate changes in output that are due to demand shocks, our estimation will correctly trace out the slope and curvature of the supply curve, regardless of whether the average markup, marginal costs, or both depend on the utilization rate.[7] We note that extensions of the model can accommodate increasing marginal costs.

Sticky prices In the presence of sticky prices, an additional mechanism generates slope and curvature in the supply curve. When prices are sticky, the capacity constraint can prevent firms from producing enough output to satisfy their demand at a given price—thus leading to a notion of rationing. As we show in Appendix C, this can pose a problem for estimating the supply curve

[7]Hence, our empirical findings below should not necessarily be interpreted as evidence for procyclical markups.

since measured prices are no longer allocative. Instead, the analogue of equation (11) has an unmeasured shadow price on the left hand side, which—in part due to rationing—is increasing and convex in the utilization rate. Since the model's prediction for how quantities respond to demand shocks remains unchanged in the presence of sticky prices, and is informative about the slope and curvature of this "shadow supply curve", we will also estimate the reduced form of the model.[8] To obtain the reduced form, we need to specify the demand side of the model and market clearing. We will do so momentarily.

Inventories Our model does not feature inventory accumulation. In practice, firms may build up inventories when demand is low and draw down inventories when demand is high. By allowing firms to decouple production from sales, inventories could reduce or eliminate the role physical capacity limits play for constraining firms' activities. This should lead us to find less or no curvature in the supply curve. In the empirical section below, we will discuss the role of inventories in detail. To foreshadow our findings, inventory holdings do not have a large effect on our estimates.

The putty-clay assumption As discussed earlier, the convexity of the supply curve hinges on the assumption that a suppliers' maximum level of output is predetermined within the period. Whether capacity can indeed be viewed as fixed depends on the time horizon and the precise definition. Throughout this paper we use the Federal Reserve's measures of capacity, which are capital-based and typically change only when the capital stock changes.[9] Adjustments in capacity therefore inherit the delays associated with capital investment. If capacity adjusted rapidly within a year (the period length in our empirical analysis), we would neither observe a large fraction of capacity constrained plants (Figure 1), nor expect to find an economically meaningful degree of curvature in the supply curve.

2.2 Demand, market clearing, and the reduced form

We next specify the demand side of the model and derive the reduced from. Estimation of the reduced form serves as a useful check since it is more robust to model misspecification and measurement problems. For instance, and as discussed above, measured prices are no longer allocative in the presence of sticky prices in our model— which could lead to misleading conclusions about the shape of the supply curve when exclusively based on the structural form. As we shall see below, estimation of the reduced form has a number of additional advantages, including that it is possible to introduce inventories in a simple way. At the same time, estimating the reduced from is not a substitute, because it does not allow us to identify the precise shape of the supply curve without taking a stance on the values of other parameters.

Since the reduced form expresses the equilibrium quantity as a function of observed and unobserved demand and supply shifters, the estimation benefits from a detailed specification of the demand side. This part of the model therefore contains sufficient detail so that industry's sales patterns match their analogues in the data.

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

[9] See Gilbert, Morin, and Raddock (2000) for details.

2.2.1 Demand and market clearing

Each industry i 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 as well as government purchases. We assume for simplicity that demand takes the constant elasticity form.

Domestic final demand Depending on whether industry i produces a consumption or investment good, private domestic final demand takes the form $C_{i,t} = \omega^C_{i,t} C_t \left(P^Y_{i,t}/P^C_t\right)^{-\sigma}$ or $I_{j,i,t} = \omega^I_{j,i,t} I_{j,t} \left(P^Y_{i,t}/P^I_{j,t}\right)^{-\sigma}$. Here, C_t are real personal consumption expenditures (PCE), and P^C_t is the PCE price index. Similarly, $I_{j,t}$ is real investment into goods of category j (e.g. equipment investment), and $P^I_{j,t}$ is the associated price index. The elasticity σ parameterizes the substitutability of varieties within each of these aggregates. Unlike quantities C_t, $I_{j,t}$ and prices P^C_t, $P^I_{j,t}$, the weights $\omega^C_{i,t}$ and $\omega^I_{j,i,t}$ are *unobserved* demand shifters.

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 $P^M_{j,t}$ the price index, its demand for industry i's output is $M_{j,i,t} = \omega^M_{j,i,t} M_{j,t} \left(P^Y_{i,t}/P^M_{j,t}\right)^{-\sigma}$. Again, $\omega^M_{j,i,t}$ is an unobserved demand shock.

Foreign demand Exports abroad constitute an additional component of industry i's demand. Demand of destination d is given by $EX_{d,i,t} = \omega^{\text{EX}}_{d,i,t} EX_{d,t} \left(P^{Y,*}_{d,i,t}/P^{\text{EX},*}_{d,t}\right)^{-\sigma}$. Here, $EX_{d,i,t}$ denotes industry i's exports to destination d, $EX_{d,t}$ is an observed demand shifter, and prices with asterisks are measured in foreign currency units. The dollar-denominated price for sales abroad is $P^Y_{i,t} = \mathcal{E}_{d,t} P^{Y,*}_{d,i,t}$, where $\mathcal{E}_{d,t}$ is the nominal exchange rate in U.S. dollars per unit of foreign currency.

Market clearing Letting $Y^{\text{inv}}_{i,t}$ denote the stock of inventories at time t, $IM_{i,t}$ imports of industry i's good, and $G_{i,t}$ sales to the government, market clearing for industry i requires that

$$Y^{\text{inv}}_{i,t-1} + Y_{i,t} + IM_{i,t} = \sum_j M_{j,i,t} + C_{i,t} + \sum_j I_{j,i,t} + G_{i,t} + \sum_d EX_{d,i,t} + Y^{\text{inv}}_{i,t}. \tag{14}$$

2.2.2 The reduced form

We next define a number of variables that capture observable shifts in demand,

$$\begin{aligned} \Delta\xi_{i,t} &= \sum_j s^M_{j,i,t-1}\Delta\ln M_{j,t} + s^C_{i,t-1}\Delta\ln C_t + \sum_j s^I_{j,i,t-1}\Delta\ln I_{j,t} + s^G_{i,t-1}\Delta\ln G_{i,t} \qquad (15)\\ &\quad + \sum_d s^{\text{EX}}_{d,i,t-1}\Delta\ln EX_{d,t} \\ \Delta\pi_{i,t} &= \sum_j s^M_{j,i,t-1}\Delta\ln P^M_{j,t} + s^C_{i,t-1}\Delta\ln P^C_t + \sum_j s^I_{j,i,t-1}\Delta\ln P^I_{j,t} + \sum_d s^{\text{EX}}_{d,i,t-1}\Delta\ln P^{\text{EX},*}_{d,t}, \qquad (16)\\ \Delta e_{i,t} &= \sum_d s^{\text{EX}}_{d,i,t-1}\Delta\ln\mathcal{E}_{d,t}. \qquad (17) \end{aligned}$$

In these expressions, $s^M_{j,i,t-1}$ denotes the sales share of industry i to downstream industry j, and other sales shares are defined similarly. $\Delta\xi_{i,t}$ is an observable demand shifter, which captures changes in the industry's customer 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 $s^M_{j,i,t-1}$ percent. Similarly, $\Delta\pi_{i,t}$ reflects changes in demand due to changes in industry i's customers' prices.

$\Delta e_{i,t}$ is the change of industry i's effective nominal exchange rate. Notice that $\Delta e_{i,t}$ varies by industry because existing trade linkages, as captured by $s^{EX}_{d,i,t-1}$, differentially expose industries to fluctuations of a common set of currencies. 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 U.S. dollars, and conditional on holding costs constant, such a depreciation leads to an increase in demand through substitution towards the industry's product.[10]

Having introduced the observable demand shifters $\Delta\xi_{i,t}$, $\Delta\pi_{i,t}$, and $\Delta e_{i,t}$, we next solve for the reduced form.

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

$$
\begin{aligned}
\Delta \ln Y_{i,t} \quad = \quad & \beta_\xi\left(\ln u_{i,t-1}\right)\Delta\xi_{i,t} + \beta_\pi\left(\ln u_{i,t-1}\right)\Delta\pi_{i,t} + \beta_e\left(\ln u_{i,t-1}\right)\Delta e_{i,t} \\
& + \beta_Q\left(\ln u_{i,t-1}\right)\Delta\ln Q_{i,t} + \beta_{mc}\left(\ln u_{i,t-1}\right)\Delta\ln mc_{i,t} \\
& + \beta_{IM}\left(\ln u_{i,t-1}\right)\frac{\Delta IM_{i,t}}{Y_{i,t-1}} + \beta_{inv}\left(\ln u_{i,t-1}\right)\frac{\Delta Y^{inv}_{i,t} - \Delta Y^{inv}_{i,t-1}}{Y_{i,t-1}} + \omega^Y_{i,t}.
\end{aligned}
\tag{18}
$$

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 $\beta_{inv} > 0$. Supply curve (11) is convex if and only if $\beta'_\xi < 0$ and $\beta'_e < 0$. The error term is a weighted average of changes in the unobserved demand shocks $\omega^C_{i,t}$, $\omega^I_{j,i,t}$, $\omega^M_{j,i,t}$, and $\omega^{EX}_{d,i,t}$.

Proof. See Appendix B. □

The equilibrium quantity is a function of all demand and supply shifters, which—in this linearized version of the model—are $\Delta\xi_{i,t}$, $\Delta\pi_{i,t}$, $\Delta e_{i,t}$, $\Delta\ln Q_{i,t}$, $\Delta\ln mc_{i,t}$, $\Delta IM_{i,t}/Y_{i,t-1}$, $\left(\Delta Y^{\text{inv}}_{i,t} - \Delta Y^{\text{inv}}_{i,t-1}\right)/Y_{i,t-1}$ as well as $\omega^Y_{i,t}$. 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 the supply curve is convex if and only if 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 3). Detailed expressions for all coefficients in Proposition 2 are listed in Appendix B.

As in Section 2.1.5 we proceed with approximating the coefficients β linearly in $\ln u_{i,t}$. If supply curves are convex, the coefficients on the interaction terms $\Delta\xi_{i,t}\cdot(u_{i,t-1} - \bar{u}_i)$ and $\Delta e_{i,t}\cdot(u_{i,t-1} - \bar{u}_i)$ will be negative, since a higher initial utilization rate leads more firms to raise markups and fewer to increase production after positive demand shocks. This intuition is illustrated in Figure (3). Before turning to the estimation, we briefly discuss a measurement problem.

[10]Note that definition (17) takes into account that some industries sell more of their goods abroad than others. If industry k sells more of its output abroad than industry ℓ, then $\sum_d s^{\text{EX}}_{d,k,t-1} > \sum_d s^{\text{EX}}_{d,\ell,t-1}$.

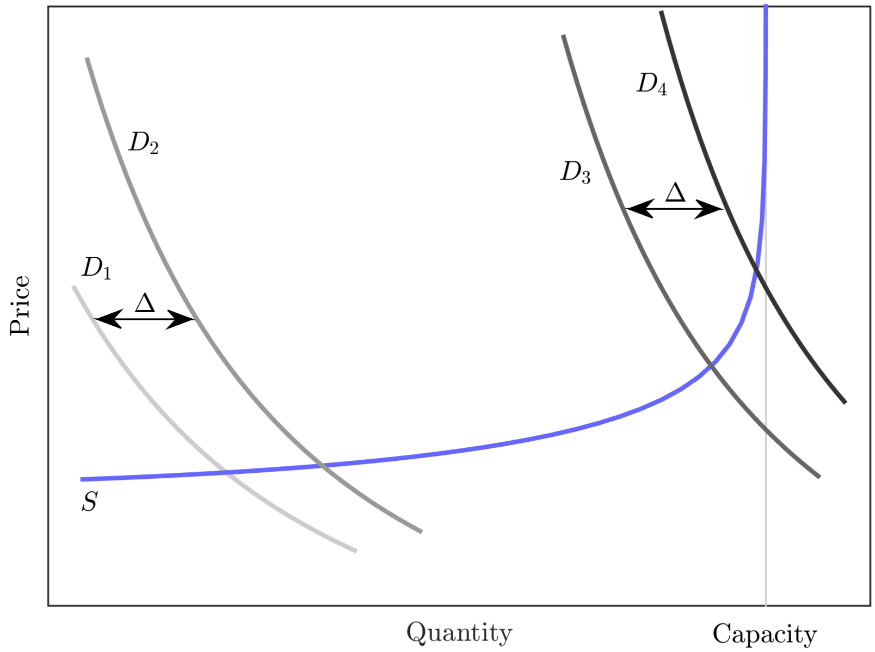


Figure 3: The utilization rate as a sufficient statistic

2.3 Measurement of marginal costs

In practice, estimation of the structural and the reduced form is complicated by the fact that marginal costs are not observed. Further, subsuming marginal costs into the error term has the two undesirable implications. First, it can lead to an omitted variable bias if the instrument is correlated with marginal costs. Second, doing so raises the variance of the estimates. It is therefore common to proxy for marginal costs with unit variable costs, which are observed. Unfortunately, this approach can also lead to biases.

In our framework industry's marginal costs differ from the industry's unit variable cost. This feature follows from the non-linear aggregation across varieties with aggregator (1). Further, the wedge between unit variable cost and marginal cost is a function of the utilization rate, that is, $\ln mc_{i,t} = \ln \mathrm{UVC}_{i,t} + \Omega(\ln u_{i,t})$, for some function Ω, where $\mathrm{UVC}_{i,t} = \left(\int_0^1 p_t^v v_t(j)\, dj\right)/Y_{i,t}$ are unit variable costs. Substituting for marginal costs in equation (11) yields

$$\ln P_{i,t}^Y = \mathcal{M}(\ln u_{i,t}) + \Omega(\ln u_{i,t}) + \ln \mathrm{UVC}_{i,t}.$$

This expression makes clear that if unit variable costs are held constant instead of marginal costs, variation in $\ln u_{i,t}$ does not identify $\mathcal{M}'$, but $\mathcal{M}' + \Omega'$, thus leading to a biased estimate. An analogous argument applies to $\mathcal{M}''$. The following proposition signs these biases.

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

Proof. See Appendix B. □

Hence, when marginal costs are proxied for with unit variable costs, estimates of the slope and curvature both exhibit a downward bias. Our estimates below should therefore be interpreted as conservative. We have verified numerically that these biases are small for reasonable calibrations, in the order of 5 to 10 percent.[11]

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

Central to the empirical analysis are the FRB's 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. Capacity is defined as the *sustainable maximum level of output, given the current capital stock* or, "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".[12] The FRB's measure of capacity is primarily based on the Survey of Plant Capacity (prior to 2007) and the Quarterly Survey of Plant Capacity (from 2007 onwards), but also uses information from the Annual Survey of Manufacturers.[13] As in our model, utilization is then calculated by dividing industrial production by capacity (see equation (10)).

Figure 4 illustrates the industry-demeaned capacity utilization rates of the 21 3-digit NAICS manufacturing industries in our sample, which enter our estimating equation (13) on the right hand side. As is clear from the figure, these 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.

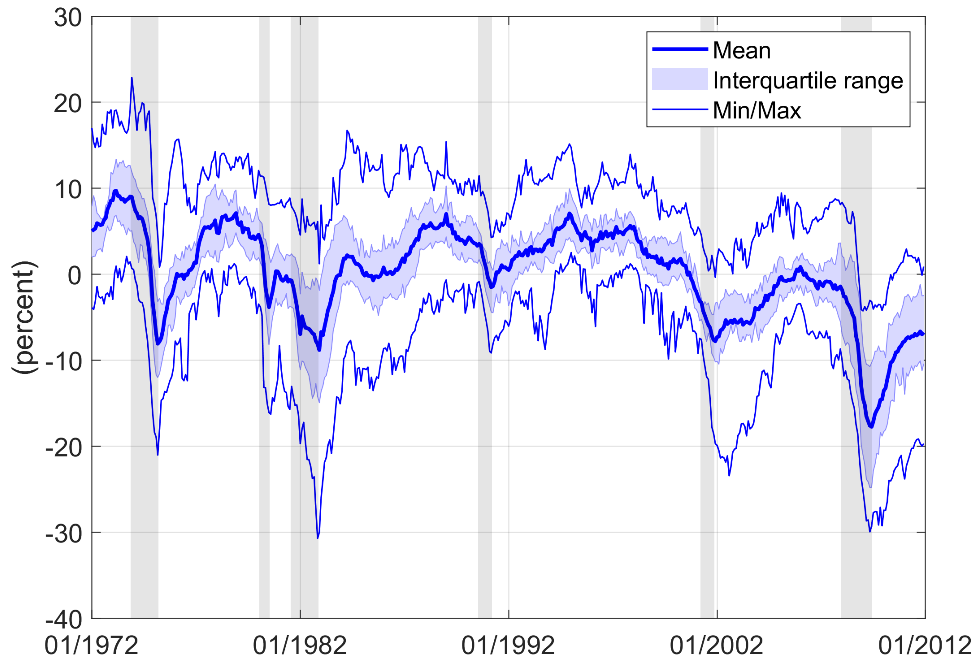
3.1.2 Additional data sources

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

[11]Regarding the intuition of Proposition 3, note that the industry's marginal costs do not depend on the utilization rate and that Ω's dependence on the utilization rate arises from the dependence of unit variable costs on the utilization rate. The sum of variable costs across plants $\int_0^1 p_t^v v_t(j)\,dj$ depends on the industry's utilization rate, and so does the aggregation of plant-level into industry-level output, equation (1). See proof of Proposition 3 in Appendix B for details. Unit variable costs are defined as the sum of plants' labor and material costs, consistent with the empirical analogue in Section 3. They do *not* reflect the aggregating firm's average costs, which include the markups of monopolistic competitors.

[12]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).

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



Notes: The figure plots the mean, minimum, maximum and interquartile range of the industry-demeaned capacity utilization series of all 21 3-digit NAICS manufacturing industries, constructed using the FRB capacity utilization data and industrial production. Shaded areas represent NBER recessions.

Figure 4: Industry-demeaned capacity utilization rates

variable costs, 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 estimates are very similar.

We calculate the sales shares $s_{j,i,t}$ 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 (15) and (16), 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 (17)) 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 $\Delta e_{i,t}$. Our sample is annual, includes all 21 3-digit NAICS manufacturing industries, and ranges from 1972 to 2011. Details on the data are available in Appendix D.

3.2 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 (13), the instrument addresses the endogeneity of $\Delta \ln Y_{i,t}$. Since we also estimate the reduced form, we use the same instrument for the demand shifter $\Delta\xi_{i,t}$ as defined in equation (15).

We consider three different instruments in our empirical analysis. In all cases, the identification assumption for the structural form 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, the controls, and the unobserved supply shocks. We emphasize that all instruments deliver comparable results.[14]

3.2.1 World import demand

The first instrument uses variation in foreign demand to estimate the supply curve (see, e.g. Hummels et al., 2014). We define the World Import Demand instrument as a weighted sum of changes in foreign output,

$$\text{inst}_{i,t}^{\text{WID}} = \sum_d s_{d,i,t-1}^{\text{EX}} \Delta \ln \text{GDP}_{d,t}. \tag{19}$$

To better understand the identifying variation, we decompose the change in foreign GDP into a common and a destination-specific component, $\Delta \ln \text{GDP}_{d,t} = \Delta \ln \text{GDP}_t^{\text{com}} + \Delta \ln \text{GDP}_{d,t}^{\text{spec}}$. 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,

$$\text{inst}_{i,t}^{\text{WID}} = \Delta \ln \text{GDP}_t^{\text{com}} \sum_d s_{d,i,t-1}^{\text{EX}} + \sum_d \bar{s}_{d,t-1}^{\text{EX}} \Delta \ln \text{GDP}_{d,t}^{\text{spec}} + \sum_d \left(s_{d,i,t-1}^{\text{EX}} - \bar{s}_{d,t-1}^{\text{EX}} \right) \Delta \ln \text{GDP}_{d,t}^{\text{spec}}. \tag{20}$$

The first term on the right hand side captures variation common to all foreign countries. Since this variation reflects the "global" business cycle (e.g. a global supply shock) and could be correlated with unobserved supply shocks, we control for it by interacting a time fixed effect with the foreign sales share of industry i, $\sum_d s_{d,i,t-1}^{\text{EX}}$. The second term on the right hand side weighs destination-specific changes in GDP with the average export share. Since our specifications will include time fixed effects—in addition to the time fixed effects interacted with the foreign sales share—this variation will be purged as well. Hence, the identifying variation of this instrument comes entirely from the third term, $\sum_d \left(s_{d,i,t-1}^{\text{EX}} - \bar{s}_{d,t-1}^{\text{EX}} \right) \Delta \ln \text{GDP}_{d,t}^{\text{spec}}$, reflecting destination-specific changes in GDP, which are weighted with the deviations of industry sales shares from the average. The identification assumption holds that this term is uncorrelated with industries' unobserved supply shocks.

[14] 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 these results.

3.2.2 Shea's instrument

Shea (1993a; 1993b) argues that output of industry j is a good instrument for output of industry i if (1) industry j demands a large share of i's output and (2) materials from i constitute a small share of j's costs. The first criterion aims at generating a high degree of relevance, while the second implies that supply shocks in industry i are unlikely to have large effects on j's output. In a simple model with input-output linkages Shea (1993b) demonstrates that the two criteria contain the degree of endogeneity, and that supply elasticities can be estimated with relatively small biases.

We extend Shea's idea by considering the possibility that shocks to industry j may affect production in industry i not only through an effect on demand. If industry i uses j's output in production, a supply shock in industry j will generally affect production in industry i also through a change in costs. To prevent this cost shock from confounding our results, we require in addition to (1) and (2) that the relationship between i and j is unidirectional. That is, we necessitate that (3) i's cost share from j is small.

We construct an instrument for industry i's output based on its sales to a select group of downstream industries, which satisfy criteria (1), (2), and (3). Recall that $s^M_{j,i,t}$ denotes industry i's sales share to industry j. Our version of Shea's instrument is then

$$\text{inst}^{\text{Shea}}_{i,t} = \sum_j s^M_{j,i,t-1} \mathbb{1}\{(1),\ (2),\ (3)\ \text{hold}\}\, \Delta \ln M_{j,t}. \tag{21}$$

The indicator function $\mathbb{1}\{(1),\ (2),\ (3)\ \text{hold}\}$ selects downstream industries j which satisfy our exogeneity criteria. We provide details on the instrument in Appendix E, but note here that, as in Shea (1993a,b), we consider both direct and indirect linkages when measuring sales and cost shares.

3.2.3 The effective exchange rate

We further use a purified change in an industry's effective exchange rate (equation (17)) to identify the slope and curvature of the supply curve. Holding costs constant, 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). Relative to the other two instruments, the effective exchange rate has the advantage that it is not predictable. We will use it to rule out that anticipation effects drive our results.

Analogous 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 $\Delta \ln \mathcal{E}_{d,t} = \Delta \ln \mathcal{E}^{com}_t + \Delta \ln \mathcal{E}^{spec}_{d,t}$. 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 (20), we decompose the effective exchange rate into three parts,

$$\Delta e_{i,t} = \Delta \ln \mathcal{E}_t^{\text{com}} \sum_d s_{d,i,t-1}^{\text{EX}} + \sum_d \bar{s}_{d,t-1}^{\text{EX}} \Delta \ln \mathcal{E}_{d,t}^{\text{spec}} + \sum_d \left(s_{d,i,t-1}^{\text{EX}} - \bar{s}_{d,t-1}^{\text{EX}} \right) \Delta \ln \mathcal{E}_{d,t}^{\text{spec}}. \tag{22}$$

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 intermediate inputs. To prevent this channel from confounding our interpretation of dollar depreciations as demand shocks, we will control for unit variable costs—which include costs of imported intermediate inputs—as suggested by the model.

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

Linear model Table 1 shows the estimates of the supply curve when we impose linearity. Specification (1) begins with Ordinary Least Squares (OLS) estimates of the inverse supply elasticity without controls. As expected, the slope estimate is insignificantly different from zero since unobserved supply shocks confound the estimation. For instance, a positive supply shock lowers prices while raising quantities, thereby biasing the slope estimate downward.

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. Controlling for unit variable costs partially addresses the simultaneity problem by removing a large fraction of the confounding variation from the error term. Specification (3) further adds the change in capacity to the equation. As predicted by the model, the coefficient is negative. All else equal, industries with greater capacity charge lower prices. That the slope coefficient rises to 0.13 suggests that the capacity control successfully purges supply shifts from the error term.

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 (equation 19), Shea's instrument (equation 21), and the effective exchange rate instrument (equation 17) in specification (5), we obtain a slope estimate of 0.23. It is greater than the OLS estimate, suggesting that despite the controls, the error term in specification (4) might still contain supply disturbances. The first stage F-statistic is 20.03. Hansen's overidentification test fails to reject the null of all instruments being valid ($p = 0.441$).

Baseline estimates We next relax the assumption of linearity and allow the inverse supply elasticity to depend on last period's utilization rate as predicted by the model. Table 2 specification (1) shows the OLS estimates. The main effect is 0.17 and the interaction term is negative and insignificant. In subsequent specifications we use instrumental variables to address the endogeneity

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

Estimator	OLS	OLS	OLS	OLS	2SLS
Instruments	-	-	-	-	WID, Shea, $\Delta e_{i,t}$
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln Y_{i,t}$	-0.09	0.08	0.13	0.17	0.23
	(0.08)	(0.02)	(0.02)	(0.02)	(0.10)
$\Delta \ln Q_{i,t}$			-0.16	-0.12	-0.16
			(0.03)	(0.04)	(0.08)
$\Delta \ln \mathrm{UVC}_{i,t}$		0.91	0.90	0.89	0.90
		(0.02)	(0.02)	(0.03)	(0.03)
R-squared	0.010	0.873	0.880	0.910	0.908
Fixed Effects†	no	no	no	yes	yes
First stage and instrument diagnostics					
F main effect					20.03
Hansen J (p-value)					0.441

Notes: The estimates are based on equation (13). Driscoll-Kraay standard errors are reported in parentheses. †: Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s_{d,i,t-1}^{\mathrm{EX}}$).

Table 1: Estimates of the linear model

of both the main effect and the interaction term, beginning with the World Import Demand instrument, Shea's instrument, the effective exchange rate, and the interaction of the World Import Demand instrument with the demeaned utilization rate $u_{i,t-1} - \bar{u}_i$. Inspection of the first stage coefficients (not reported) suggests that, by and large, the first three of these instruments explain $\Delta \ln Y_{i,t}$, while the interaction of the World Import Demand instrument with the utilization rate explains the interaction $\Delta \ln Y_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$. As specification (2) in Table 2 shows, the coefficient on the interaction term becomes positive and significant. That is, the slope of the supply curve is increasing in the initial capacity utilization rate, implying that supply curves are convex.

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 in 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 1.13 and precisely estimated. These estimates are reported in specification (5).

The size of the interaction term is economically sizable. Since the utilization rate is demeaned, the estimated main effect of specification (5) implies that, all else equal, an increase in output by one percent raises prices on average by 0.26 percent. Raising the initial utilization rate by two standard deviations ($2 \cdot 0.066$) raises the inverse supply elasticity to 0.41 ($= 0.26 + 2 \cdot 0.066 \cdot 1.13$). We discuss the quantitative implications of our preferred estimates below.

To assess instrument relevance in the presence of multiple endogenous regressors, we report

Dependent variable: $\Delta \ln P^{Y}_{i,t}$

Estimator	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Instruments:						
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)
$\Delta \ln Y_{i,t}$	0.17	0.27	0.28	0.26	0.27	0.26
	(0.02)	(0.09)	(0.09)	(0.10)	(0.08)	(0.09)
$\Delta \ln Y_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-0.33	0.97	1.36	0.92	1.13	1.13
	(0.24)	(0.31)	(0.72)	(0.66)	(0.33)	(0.33)
$u_{i,t-1} - \bar{u}_i$	0.01	0.03	0.03	0.03	0.03	0.03
	(0.02)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
$\Delta \ln Q_{i,t}$	-0.11	-0.23	-0.24	-0.21	-0.22	-0.21
	(0.04)	(0.11)	(0.11)	(0.12)	(0.10)	(0.11)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-0.32	-1.13	-1.35	-1.09	-1.21	-1.20
	(0.43)	(0.42)	(0.50)	(0.43)	(0.42)	(0.42)
$\Delta \ln \mathrm{UVC}_{i,t}$	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	0.902	0.897	0.903	0.901	0.901
Fixed Effects	yes	yes	yes	yes	yes	yes
First stage and instrument diagnostics†						
F Main effect		10.84	19.23	12.05	18.00	17.27
F Interaction		16.84	6.43	3.82	65.87	28.25
Cragg-Donald Wald F		9.08	7.76	6.67	6.22	9.02
SW F Main effect		31.30	20.21	25.65	17.97	31.42
SW F Interaction		33.76	16.86	6.01	25.10	33.52
Hansen J (p-value)		0.409	0.380	0.407	0.719	0.531

Notes: The estimates are based on equation (13). Driscoll-Kraay standard errors are reported in parentheses. Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s^{\mathrm{EX}}_{d,i,t-1}$).

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table 2: Estimates of the non-linear model

the Sanderson and Windmeijer (2016) conditional (SW) F-statistic and the Cragg-Donald Wald F statistic in addition to the standard F-statistics in the bottom panel of Table 2. These diagnostics suggest that the effective exchange rate instrument is potentially weak. When we drop it and its interaction term with the utilization rate from our set of instruments, the estimated coefficients both remain positive and highly significant (specification (6)). Under the identification assumption that conditional on controls, the WID and Shea's instrument are orthogonal to unobserved supply shocks, supply curves at the industry level are increasing and convex. Further, all diagnostics suggest that this set of instruments is strong.[15]

Ad-hoc estimation of the supply curve In Appendix F.1 we report results from an ad-hoc estimation of the supply curve. In particular, we regress $\Delta \ln P^Y_{i,t}$ on a linear and a squared term in $\Delta \ln Y_{i,t}$, using the WID and Shea's instrument as well as their squares to address simultaneity. As Appendix Table F1 shows, the first stage for $(\Delta \ln Y_{i,t})^2$ is always weak ($F < 3$). This contrasts to the first stage of the interaction term as reported in Table 2, and highlights the empirical usefulness of the model: With our instruments, it is not possible to estimate the curvature of the supply curve with a squared term. We can only estimate the curvature, when imposing the structure of the model.

3.3.2 Robustness and extensions

We next turn to a number of robustness checks and extensions of the baseline analysis.

Inventories As noted earlier, by holding sufficient inventories, firms may decouple production from sales and thereby reduce the degree to which capacity constraints limit production. Our estimates this far indicate that capacity constraints generate convex supply curves *despite* the possible presence of inventories. We next explore the role of inventories in greater detail.

Figure 5 shows a histogram of industry's inventory holdings relative to annual production. The histogram is based on the same sample as the estimation. For most industry-year observations, inventories cover between one and two months worth of production. They rarely exceed three months worth of production. Under the assumption that the goods held in stock match those in higher demand, the figure suggests that industries may temporarily satisfy higher demand by running down their inventories.

Table 3 shows correlations of the level and changes of inventory holdings with the utilization rate. With a correlation coefficient of 0.02 the utilization rate and the level of inventory holdings are essentially uncorrelated. The correlation of the utilization rate at time t with the change in inventories from $t-1$ to t is 0.43. Further, the correlation with the change in inventories from t to $t+1$ is 0.12. These positive correlations indicate that industries tend to *increase* their inventories—rather than running them down—when the utilization rate is high.[16] It is therefore not clear whether firms in practice use inventories to escape capacity constraints.

We next turn to a number of robustness checks of our regression analysis. The starting point for these checks is specification (6) of Table 2, which estimates the slope and curvature of the supply curve using the World Import Demand instrument, Shea's instrument, and their interactions with the utilization rate. In specification (1) of Table (4), we include the industry-demeaned lag of inventory holdings ($Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$) as a control. In specification (2) we additionally

[15] For the standard F statistic and the SW conditional F statistic, we apply the conventional threshold of 10. That is, an instrument is potentially weak if the F statistic is below 10. Stock and Yogo (2005) tabulate critical values for weak IV tests based on the Cragg-Donald statistic. These critical values depend on the number of endogenous right hand side variables and the number of instruments. In specification (6) of Table 2, a value of 9.02 implies that the maximum bias of the 2SLS estimator is less than 10 percent of the bias of the OLS estimator (the critical value for a maximal 10 percent relative bias is 7.56). We have also checked that our estimates are robust to using a limited information maximum likelihood estimator (LIML), which has better small sample properties than 2SLS in the presence of weak instruments. The estimates of specification (6) in Table 2 with the LIML estimator are almost identical.

[16] That inventory investment is procyclical is well documented. Bils and Kahn (2000) offer an explanation based on the assumption that inventories facilitate sales.

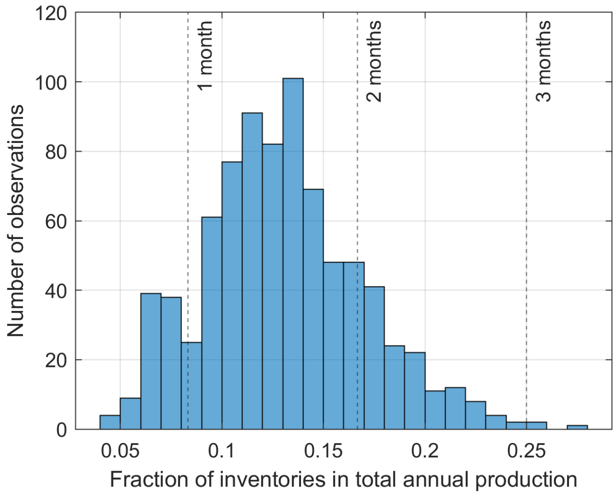


Figure 5: Histogram of inventory holdings

Correlations				
	$u_{i,t}$	$\frac{Y^{\text{inv}}_{i,t}}{Y_{i,t}}$	$\frac{\Delta Y^{\text{inv}}_{i,t}}{Y_{i,t-1}}$	$\frac{\Delta Y^{\text{inv}}_{i,t+1}}{Y_{i,t}}$
$u_{i,t}$	1.00			
$\frac{Y^{\text{inv}}_{i,t}}{Y_{i,t}}$	0.02	1.00		
$\frac{\Delta Y^{\text{inv}}_{i,t}}{Y_{i,t-1}}$	0.43	0.07	1.00	
$\frac{\Delta Y^{\text{inv}}_{i,t+1}}{Y_{i,t}}$	0.12	-0.06	0.10	1.00

Note: $u_{i,t}$ and $\frac{Y^{\text{inv}}_{i,t}}{Y_{i,t}}$ are industry-demeaned.

Table 3: Utilization rates and inventories

include an interaction term with the change in output. When doing so, we add interactions of the World Import Demand instrument and Shea's instrument with $Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$ to the set of instruments. The objective of both specifications is to trace out the slope and curvature of the supply curve, holding the initial level of inventories constant. In both cases the estimates remain essentially unchanged. In specification (3) we further include the contemporaneous and lagged change in inventories as controls. Again, the estimates are barely effected.

In specification (4) of Table (4) we drop from the sample observations for which the initial level of inventories exceeds two months worth of production. In this "low-inventory" sample, the slope and curvature estimates are higher than in the full sample (specification (6) of Table 2). While this finding is qualitatively consistent with the view that inventories allow firms to reduce the frequency of hitting their capacity constraints, the effect on the estimates is quantitatively small. Further, when we instead drop observations with less than one month worth of inventories in specification (5), the slope and curvature also rise relative to the baseline (though less than in the low-inventory sample).[17] In summary, neither of these checks indicates that the presence of inventories strongly affects our results.

The unit cost control As discussed in Section 2.3, proxying marginal costs with unit variable costs may lead to downward-biased estimates of the slope and curvature in our model. Consistent with this prediction and as shown in Table 5 specification (1), the estimates of both increase, when we instead subsume marginal cost changes into the error term. While these estimates likely exhibit less bias, the error bands also increase substantially. We therefore prefer to include unit variable costs as a control and to interpret our estimates as conservative.

[17] When we split the sample in the middle, the instruments are weak in both subsamples.

Dependent variable: $\Delta \ln P^{Y}_{i,t}$

	(1)	(2)	(3)	(4)	(5)
Sample	full	full	full	$Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} \leq 2/12$	$Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} \geq 1/12$
$\Delta \ln Y_{i,t}$	0.26	0.27	0.27	0.31	0.30
	(0.09)	(0.08)	(0.09)	(0.11)	(0.09)
$\Delta \ln Y_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	1.13	1.12	1.18	1.55	1.26
	(0.33)	(0.31)	(0.30)	(0.48)	(0.31)
$u_{i,t-1} - \bar{u}_i$	0.02	0.03	0.06	0.02	0.06
	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)
$\Delta \ln Q_{i,t}$	-0.21	-0.22	-0.22	-0.27	-0.29
	(0.11)	(0.10)	(0.09)	(0.14)	(0.11)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-1.19	-1.16	-1.29	-1.35	-1.28
	(0.42)	(0.48)	(0.44)	(0.62)	(0.48)
$\Delta \ln \text{UVC}_{i,t}$	0.89	0.89	0.90	0.91	0.82
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
$Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$	-0.04	-0.03	0.03		
	(0.08)	(0.08)	(0.09)		
$\Delta \ln Y_{i,t} \cdot \left(Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}\right)$		-0.33	-0.41		
		(1.05)	(1.10)		
$\Delta Y^{\text{inv}}_{i,t}/Y_{i,t-1}$			0.15		
			(0.17)		
$\Delta Y^{\text{inv}}_{i,t-1}/Y_{i,t-2}$			-0.44		
			(0.15)		
R-squared	0.901	0.900	0.903	0.900	0.872
Fixed Effects	yes	yes	yes	yes	yes
Observations	819	819	819	673	719
First stage and instrument diagnostics†					
F Main effect	18.16	21.97	17.14	11.56	14.00
F Interaction w/ $u_{i,t-1} - \bar{u}_i$	28.41	21.83	22.00	21.04	22.77
F Interaction w/ $Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$		50.37	62.17		
Cragg-Donald Wald F	8.98	6.16	6.45	7.33	7.27
SW F Main effect	33.05	30.48	17.85	9.58	29.58
SW F Interaction w/ $u_{i,t-1} - \bar{u}_i$	33.69	34.00	38.24	17.79	33.16
SW F Interaction w/ $Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$		88.72	81.56		
Hansen J (p-value)	0.522	0.516	0.346	0.576	0.515

Notes: The estimates are based on equation (13) using 2SLS with the World Import Demand and Shea's instrument as well as interactions of both with $u_{i,t-1} - \bar{u}_i$. The set of instruments for specifications (2) and (3) further includes interactions of the World Import Demand and Shea's instrument with $Y^{\text{inv}}_{i,t-1}/Y_{i,t-1} - \overline{Y^{\text{inv}}_i/Y_i}$. Driscoll-Kraay standard errors are reported in parentheses. Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s^{\text{EX}}_{d,i,t-1}$).

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table 4: The role of inventories

The capacity control and anticipation effects In specifications (2) and (3) of Table 5 we examine the implications of dropping the change in capacity and its interaction with the utilization rate from the regression. Consistent with the model's prediction that changes in capacity shift the supply curve, the estimates of the slope and curvature fall (relative to specification (6) of Table 2). Hence, the change in capacity is a useful control variable—even when the supply curve is estimated with instrumental variables.

If the model is correctly specified anticipation effects do not pose a problem for identification, because the observed change in capacity captures all relevant information about future shocks (see Section 2.1.6). To the extent that the model is incorrectly specified, anticipation effects could still pose a problem for identification. We address this concern in the next section by estimating the effect of exchange rate changes on output (the reduced form). Since the size of this effect decreases with the initial utilization rate (equation (18)), and changes in the exchange rate are not predictable, anticipation effects do not drive the estimated curvature of the supply curve.

Sticky prices We next add a number of additional controls to the equation. Specification (4) of Table 5 adds the percent change of the industry's price from t to $t+1$. Extensions of the model to include sticky prices suggest that this variable should capture the firm's expectations about changes in future marginal costs (see Appendix C). Adding this control has virtually no effect on the estimates of slope and curvature. Further, the coefficient on future price changes is close to zero and tightly estimated, suggesting that producer prices are flexible over a year-long horizon.[18]

Higher order terms Our estimating equation (12) is a first order approximation around $t-1$ values. Relative to an approximation around the steady state, the approximation around $t-1$ values allows us to estimate the curvature of the supply curve with an interaction term (see discussion in Section 2.1.5 and Figure 3). It is natural to ask whether our results are robust to the inclusion of higher order terms as controls.

As noted above, the inclusion of endogenous second order terms on the right hand side leads to a weak instrument problem. Instead, we include an interaction term of changes in unit variable costs with the utilization rate in specification (5). This variable is positive and highly significant, suggesting that pass-through of cost shocks into prices is stronger when capacity utilization is high. However, the slope and curvature of the supply curve change little. To better control for industries' "initial position" on their supply curve, specification (6) includes the square of the lagged utilization rate as a control. Doing so raises both the slope and the curvature estimates. Adding a squared term of the change in capacity as a control has no meaningful impact on the estimated curvature (estimates not reported).

Further robustness In specification (7) of Table 5, we include $\Delta\pi_{i,t}$ as defined in equation (16) as a control to capture unobserved inflationary pressure downstream. The estimates change little. Specification (8) includes all controls simultaneously. In this preferred specification the main effect on output is 0.33 and the coefficient on the interaction term is 1.51. We plot the associated inverse supply elasticity is in Panel A of Figure 6 below.

[18] A second diagnostic that suggests that producer prices are quite flexible when differenced over one year is the high pass-through of unit variable costs changes into price changes. In models with sticky prices, this pass-through is substantially less than one. Our estimates suggest that it is close to 0.9.

One potential concern with the estimates is that there is a purely mechanical correlation between the price change on the left hand side and the change in unit variable costs on the right hand side. In all specifications this far, the price index has been constructed as an implicit deflator by dividing the market value of production by the index of industrial production, and the unit variable cost measure on the right hand side was constructed by dividing variable costs by the index of industrial production. The common division by industrial production could therefore induce a purely mechanical correlation. As specification (1) in Table 5 showed, the slope and curvature estimates are not driven by this correlation and increase when unit variable costs are dropped from the regression. As an additional check, we use the price index from the NBER-CES manufacturing industry database instead of our preferred implicit price measure on the left hand side. The estimates, reported in specification (9) of Table 5, are very similar.

Coefficient restrictions The model implies a number of coefficient restrictions for estimating equation (12). Among others, the model predicts that the slope coefficient of the supply curve is the negative of the coefficient on capacity (and similarly for the interaction terms). Inspection of the estimates in specification (6) of Table 2 suggests that these cross-coefficient restrictions broadly hold. In Appendix F.2 we test these restrictions formally. The model performs well.

Nonparametric estimation of the structural form Figure 6 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 align well with those based on the interaction term.

The figure also demonstrates that at low levels of the utilization rate, the estimated inverse supply elasticity is statistically indistinguishable from zero. This contrasts to a value of 0.48 at the 95th percentile ($u_{i,t-1} - \bar{u}_i = 0.10$), which is highly statistically significant. The non-parametric estimate for the highest utilization rates suggests an even larger value of 0.61. Comparing observations with utilization rates below the 15th and above the 85th percentile, the non-parametric estimates imply that a one percent outward shift in demand raises prices by approximately 0.1 percent in the former and by 0.6 percent in the latter case (a sixfold difference). Panel B of Figure 6 plots the quadratic and the partially-linear approximation of the supply curve. In the spirit of Nakamura and Steinsson (2018) we note that these estimates can be used as “identified moments” in other models, including in models without a concept of capacity.

Heterogeneity We discuss cross-industry heterogeneity of supply curves in Appendix F.3. For instance, we show that there are no statistically significant differences in the shapes of supply curves of durable and nondurable goods producing industries. For both groups of industries supply curves slope up and are convex.

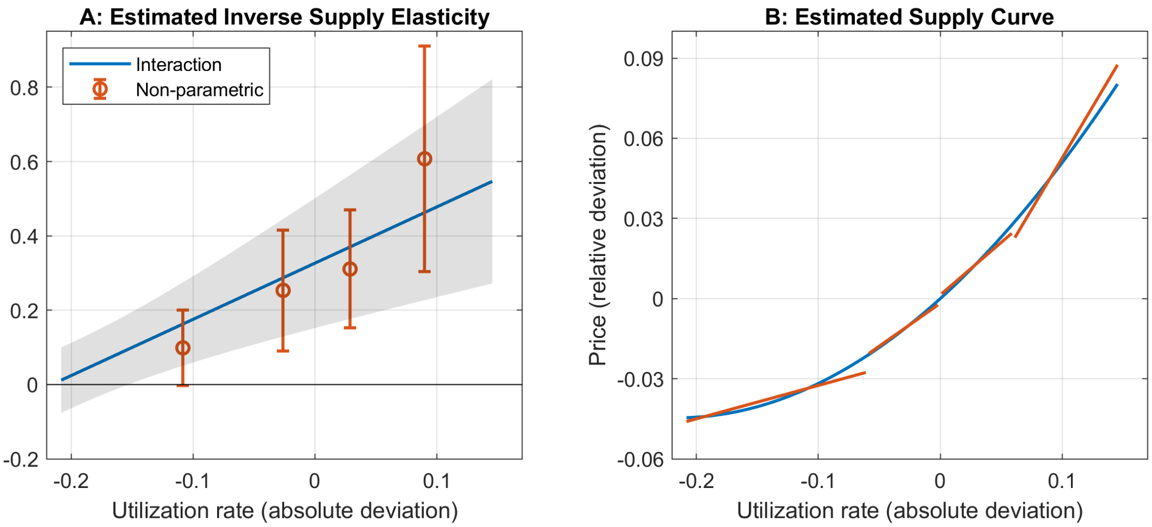
Dependent variable: $\Delta \ln P^Y_{i,t}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta \ln Y_{i,t}$	0.69	0.24	0.22	0.26	0.26	0.35	0.26	0.33	0.28
	(0.28)	(0.09)	(0.07)	(0.09)	(0.08)	(0.10)	(0.10)	(0.11)	(0.06)
$\Delta \ln Y_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	1.85	0.95	0.82	1.13	0.99	1.70	1.21	1.51	1.15
	(0.91)	(0.32)	(0.27)	(0.34)	(0.31)	(0.44)	(0.34)	(0.46)	(0.23)
$u_{i,t-1} - \bar{u}_i$	0.44	0.02	-0.02	0.03	0.00	0.09	0.03	0.06	0.00
	(0.14)	(0.04)	(0.02)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.02)
$\Delta \ln Q_{i,t}$	-0.80	-0.19		-0.21	-0.20	-0.31	-0.22	-0.28	-0.17
	(0.32)	(0.11)		(0.11)	(0.10)	(0.11)	(0.12)	(0.12)	(0.06)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-2.29			-1.21	-1.03	-2.90	-1.38	-2.44	-1.75
	(1.13)			(0.43)	(0.41)	(0.68)	(0.41)	(0.85)	(0.54)
$\Delta \ln \mathrm{UVC}_{i,t}$		0.89	0.89	0.89	0.90	0.90	0.85	0.87	0.92
		(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)	(0.08)
$\Delta \ln P^Y_{i,t+1}$				0.00				0.01	-0.11
				(0.02)				(0.02)	(0.06)
$\Delta \ln \mathrm{UVC}_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$					1.13			0.74	1.72
					(0.29)			(0.35)	(0.47)
$(u_{i,t-1} - \bar{u}_i)^2$						1.25		0.90	0.55
						(0.30)		(0.33)	(0.23)
$\Delta \pi_{i,t}$							0.27	0.25	-0.09
							(0.09)	(0.10)	(0.14)
R-squared	0.469	0.902	0.901	0.901	0.907	0.896	0.902	0.903	0.855
Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
First stage and instrument diagnostics†									
F Main effect	16.42	17.14	18.10	16.12	17.51	16.20	17.51	15.76	11.14
F Interaction	27.80	28.97	31.58	27.58	23.72	10.73	23.55	8.70	15.96
Cragg-Donald Wald F	8.68	9.29	9.97	8.66	9.09	7.33	9.03	7.23	5.06
SW F Main effect	29.08	33.25	21.80	28.03	26.93	15.79	31.90	14.01	15.81
SW F Interaction	31.64	38.72	35.35	30.66	29.72	20.03	30.73	16.30	34.85
Hansen J (p-value)	0.297	0.484	0.484	0.529	0.748	0.731	0.524	0.629	0.348

Notes: The 2SLS estimates are based on equation (13) using the World Import Demand and Shea's instrument as well as interactions of both with $u_{i,t-1} - \bar{u}_i$. Driscoll-Kraay standard errors are reported in parentheses. Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s^{\mathrm{EX}}_{d,i,t-1}$).

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table 5: Robustness of the non-linear model



Notes: The bins of the utilization rate $u_{i,t-1} - \bar{u}_i$ are 1) below -0.06 (approximately the 15th percentile), 2) between -0.06 and 0, 3) between 0 and 0.06, and 4) above 0.06 (approximately the 85th percentile). The parametric estimates are based on specification (8) of Table (5).

Figure 6: Non-parametric estimates

3.3.3 Estimates of the reduced form

We next turn to the estimates of the reduced form (equation 18). To conserve on space we discuss the linear estimates of the reduced form in Appendix F.4. Specification (1) in Table 6 presents OLS estimates of the non-linear reduced form. This specification is based on equation (18) and all coefficients are allowed to depend linearly on the utilization rate. The coefficient on the interaction term $\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ is negative and significant. That demand shocks stimulate production more when the utilization rate is initially low implies that supply curves are convex (Proposition 2).

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. This estimate again implies that supply curves are convex. Importantly, it also implies that anticipation effects do not drive our curvature estimates, since changes in exchange rates are not predictable.

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 -4.49, and with Shea's instrument it becomes -2.89 (specification (3)). When we combine both instrument in specification (4), the coefficient is -3.57. Hence, the initial utilization rate robustly determines how much production responds to demand shocks. The instruments are uniformly strong. In Appendix F.4 we discuss robustness of the reduced form estimates and in Appendix F.5 we show that the estimates of the structural and reduced forms are broadly consistent with one another.

Nonparametric estimates Figure 7 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

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

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)
$\Delta\xi_{i,t}$	0.75	0.82	0.83	0.82
	(0.09)	(0.14)	(0.13)	(0.13)
$\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-1.78	-4.49	-2.89	-3.57
	(0.61)	(1.02)	(1.59)	(0.95)
$\Delta e_{i,t}$	0.50	0.30	0.36	0.34
	(1.02)	(0.90)	(0.91)	(0.90)
$\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-20.28	-20.76	-20.61	-20.67
	(5.30)	(5.93)	(5.45)	(5.61)
$u_{i,t-1} - \bar{u}_i$	-0.23	-0.19	-0.21	-0.20
	(0.06)	(0.04)	(0.05)	(0.04)
R-squared	0.846	0.840	0.845	0.843
Other controls	yes	yes	yes	yes
Fixed Effects	yes	yes	yes	yes
First stage and instrument diagnostics†				
F Main effect		115.26	140.27	109.54
F Interaction		14.33	16.12	18.31
Cragg-Donald Wald F		28.77	39.33	38.28
SW F Main effect		173.66	215.30	149.34
SW F Interaction		21.44	24.64	25.47
Hansen J (p-value)		0.447	0.442	0.559

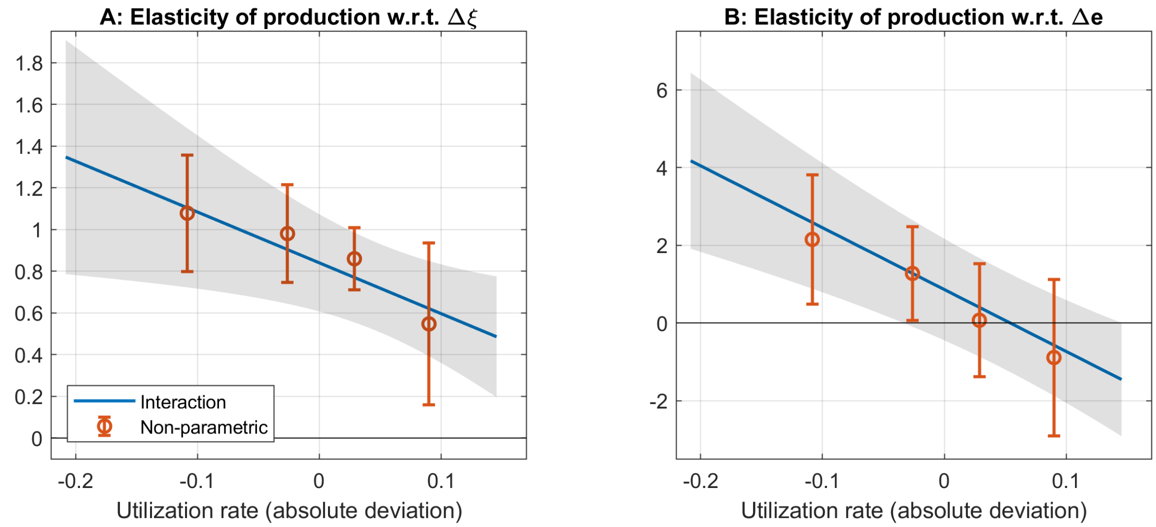
Notes: The estimates are based on equation (13). Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s^{\text{EX}}_{d,i,t-1}$). Other controls include $\Delta\pi_{i,t}$, $\Delta \ln Q_{i,t}$, $\Delta \ln \text{UVC}_{i,t}$, $\Delta IM_{i,t}/Y_{i,t-1}$, $\left(\Delta Y^{\text{inv}}_{i,t} - \Delta Y^{\text{inv}}_{i,t-1}\right)/Y_{i,t-1}$, and their interactions with $u_{i,t-1} - \bar{u}_i$. Driscoll-Kraay standard errors are reported in parentheses.

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table 6: Estimates of the non-linear reduced form

percentile (0.10). The exchange rate response drops to zero at high utilization rates (Panel B), although we note that the main effect is imprecisely estimated (Table 6).

Summary of results In summary, our estimates indicate a sizable degree of curvature of the supply curve. Price responses to increases in demand are much larger at high levels of capacity utilization, and production responses smaller. The evidence for convex supply curves is robust across specifications and independent of whether we estimate the structural or the reduced form.



Notes: The bins of the utilization rate $u_{i,t-1} - \bar{u}_i$ are 1) below -0.06 (approximately the 15th percentile), 2) between -0.06 and 0, 3) between 0 and 0.06, and 4) above 0.06 (approximately the 85th percentile). The parametric estimates are based on specification (6) of Table F5.

Figure 7: Non-parametric estimates of the reduced form

4 General equilibrium applications

The objective of this section is to study the aggregate implications of supply curves that are convex at the industry level. To do so, we embed a variant of the partial equilibrium framework from Section 2 in a general equilibrium environment with input-output linkages. Although convex supply curves have implications in a variety of other contexts, we focus for concreteness on the following three applications. First, we study the state-dependence of the aggregate government expenditure multiplier. Second, we study the mapping from output (Y) to real marginal costs (mc/P^Y), an integral component of the slope of the Phillips curve. This mapping becomes convex in our framework. Third, we calculate the welfare costs of business cycles in our model. These are larger than in standard frameworks such as Lucas (1987).

4.1 Model setup and calibration

At the heart of the model is a set of industries, which jointly form a nonlinear production network. Since our estimates in Section 3 are based on a sample of manufacturing industries, it is not clear to what extent they are valid for non-manufacturing industries. We therefore model two different types of industries. In some industries firms produce subject to a capacity constraint. We will refer to these industries as capacity industries. The remaining *non-capacity* industries have a representative firm operating a standard neoclassical production technology with constant returns to scale. To focus our attention on the implications of convex supply curves, the remainder of the model is standard. In contrast to Section 2.2, the economy is closed.

Model setup A representative household consumes and supplies labor to maximize utility subject to her budget constraint. Labor is specific to each industry. The household purchases a consumption bundle composed of the goods and services produced by 71 industries. Some of these industries exhibit a notion of capacity while others do not. Capacity industries are modelled similar to those in Section 2, but with constant productivity and a fixed capital stock. Non-capacity industries use a technology that is linear in variable inputs. The variable input bundle includes labor and a bundle of materials, where the relative weights on labor and materials from other industries are calibrated to the data. The business cycle is driven by a shock that scales the marginal utility of consumption, which we interpret as a demand shock. We provide a detailed description of the model in Appendix G.

Calibration We briefly highlight key elements of the calibration. Of the 71 industries in the model, our benchmark calibration assumes that 45 exhibit a concept of capacity and the remaining 26 do not. This classification is based on whether an industry's output is likely to be constrained by capacity when demand increases while the capital stock is held fixed. For instance, we classify construction as a capacity industry, but not retail or wholesale industries. We document our choices in Appendix G.3. Elasticity θ and the distribution of idiosyncratic demand shocks (see Section 2) are calibrated so that the supply curve of capacity industries matches our preferred estimate in the data. As Appendix Figure G1 shows, the fit is almost exact. We further choose the weights in industries' production functions so that their cost shares match those in the BEA's Use Tables. Two additional parameters are important for our results. First, we choose a Frisch labor supply elasticity of 2. This value is greater than most micro estimates, but a common choice in DSGE models and in line with the argument in Hall (2009). Second, based on a number of recent estimates, we assume that the substitutability of inputs is very limited in the short run and set the elasticities of substitution to 0.05 in the benchmark calibration. We discuss this choice momentarily. Details on the calibration are available in Appendix G.3.

Related literature Our framework is related to prior work studying the transmission of shocks in non-linear production networks. Baqaee and Farhi (2019) emphasize the aggregate implications of nonlinearities in production networks for applications such as the asymmetric distribution of GDP growth rates and the welfare costs of business cycles.[19] Several other papers study the role of sectoral shocks for the aggregate business cycle, including Bigio and La'O (2016). Our approach differs from this literature, because we explicitly model capacity industries and carefully estimate them in the data. Further, we are interested in aggregate demand shocks rather than industry-specific productivity distortions.

4.2 Mechanisms that affect aggregate curvature

Three mechanisms determine the overall curvature of the aggregate supply curve and hence aggregate outcomes. First, it is known that markups accumulate in production networks. If one industry sells its output to another industry before the good is sold to the final consumer, the price that the consumer pays includes the markups from both industries. Put differently, net markups are additive in production networks. In the context of our model, this additivity of markups constitutes a powerful mechanism that leads the aggregate supply curve to exhibit greater slope and curvature than industries' supply curves.

[19] Note that Hulten's (1978) theorem in its original form does not hold in our framework, because labor supply is endogenous and firms set time-varying markups.

On the other hand, substitution attenuates the aggregate slope and curvature relative to the industry level. If the relative price of an industry rises, firms and households substitute away from its product. While substitution is likely important for similar goods and over long time horizons, our benchmark calibration follows a recent empirical literature that documents that substitutability is low for broad industry categories and at business cycle frequencies (e.g. Atalay, 2017, Boehm, Flaaen, and Pandalai-Nayar, 2019). We also consider a case with a higher elasticity of substitution below.

Third, if aggregate shocks affect industries differentially, this heterogeneity affects the slope and curvature of the aggregate supply curve. In the presence of nonlinearities—and holding all else equal—a shock that lowers two otherwise identical industries' output by 5 percent, has a different aggregate effect than a shock which reduces one industry's output by, say, 2 percent and the other's by 8 percent. In the latter case, the effect of industries' convex supply curves on aggregate curvature is larger. It turns out that this heterogeneity channel is quantitatively unimportant in our model, because there is no mechanism (other than substitution) that generates differential responsiveness to aggregate shocks. We discuss details in Appendix G.5.

4.3 Applications

We next turn to three applications.

4.3.1 The government expenditure multiplier

Convex supply curves imply that the government expenditure multiplier is countercyclical. Estimates of the multiplier for different states of the business cycle vary and the literature has not yet converged to a consensus (see Section 1). This is, in part, because aggregate time series variation is too limited to deliver precise estimates. Our contribution here is to quantify the role of convex supply curves that are precisely estimated at the industry level. This quantification proceeds as follows. We first choose the demand shock to generate an output gap relative to the steady state. For each value of the output gap, we then compute the fiscal multiplier for an instantaneous and small government expenditure shock as dY/dG.

As Table 7 shows, the multiplier for the baseline calibration ranges from 0.55 to 0.68 for output gaps between negative and positive 10 percent. Hence, the multiplier is approximately 12 cents on the dollar greater when output is 10 percent below potential, than when output is 10 percent above potential. As the table shows, this difference depends on the calibration. It falls for greater elasticities of substitution, and rises if more industries are capacity industries or when the Frisch labor supply elasticity is greater. While these differences in multipliers are modest relative to the estimates in Auerbach and Gorodnichenko (2012, 2013), we note that our model contains no mechanisms other than convex supply curves that affect the state-dependence of the multiplier. A similar mechanism in the labor market would amplify this degree of state-dependence (Michaillat, 2014).

4.3.2 The mapping from output to real marginal costs

We next turn to the mapping from output to real marginal costs. We focus on this mapping because it is well defined even in static models and plays a key role for the slope and curvature

	Initial output relative to steady state (in percent)					Difference
	-10	-5	0	5	10	-10/+10
Calibration	Government expenditure multiplier					
Baseline	0.68	0.65	0.63	0.59	0.55	0.12
Elasticity of substitution 0.5	0.68	0.66	0.64	0.63	0.61	0.07
22 capacity industries	0.72	0.71	0.69	0.67	0.64	0.08
66 capacity industries	0.62	0.59	0.56	0.51	0.46	0.16
Frisch labor supply elasticity 1	0.56	0.54	0.52	0.50	0.47	0.09
Frisch labor supply elasticity 5	0.77	0.74	0.71	0.67	0.62	0.15

Notes: The baseline calibration has 45 capacity industries, an elasticity of substitution of 0.05, and a Frisch labor supply elasticity of 2. See Appendix G.3 for details. In the BEA's industry classification 22 industries are in the manufacturing sector, and 66 industries are private.

Table 7: State-dependent multipliers

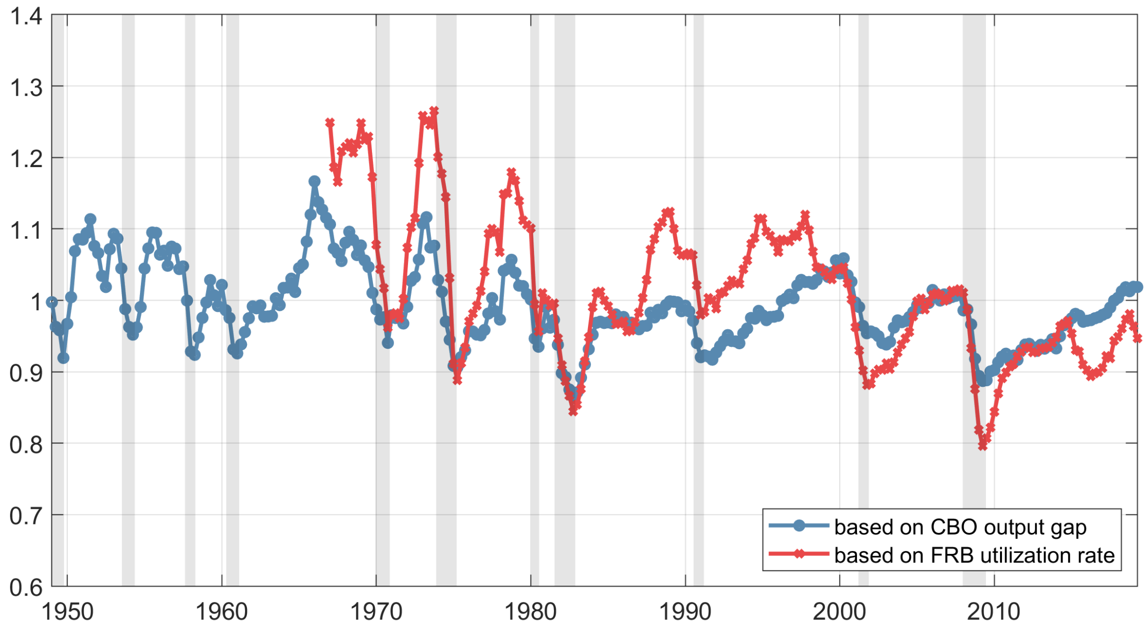
of the Phillips curve. To be precise, we define the partial elasticity of real marginal costs with respect to output as

$$\left.\frac{\partial \ln \frac{mc_t}{P_t^Y}}{\partial \ln Y_t}\right|_{MU=const}.$$

Calculation of this elasticity requires the entire model as described in Appendix G.1, except that 1) we do not impose output market clearing so that output can be varied independently, 2) we do not impose that intermediaries sell their output to the final goods sector at constant markups (as in New Keynesian models), and 3) we hold the marginal utility of consumption constant. We define this elasticity more rigorously in Appendix G.4.

As Figure 8 shows, a convex aggregate supply curve implies that the partial elasticity of real marginal costs with respect to output is procyclical. In standard sticky price models, this leads to a flattening of the Phillips curve during downturns. If a positive demand shock hits the economy during a downturn, this shock generates less inflationary pressure than if the same-sized shock hits during an expansion. For instance, during the Great Recession, the slope of the Phillips curve falls by approximately 12 to 20 percent. This finding can partially explain the "missing disinflation" observed at the time, and has implications for stabilization policy that relies on generating inflation. If episodes at the zero lower bound coincide with low output, government expenditures will generate less inflation than predicted by models in which the Phillips curve is log-linear (e.g. Christiano, Eichenbaum, and Rebelo, 2011).

Note also that when we compute the partial elasticity of real marginal costs with respect to output using the demeaned utilization rate from the FRB—the red line in Figure 8—it trends downwards. The reason is that the utilization rate itself trends downwards. When taken at face value, this implies that the Phillips curve has significantly flattened since the 1970's, although the downward trend in the utilization rate raises a number of questions (Pierce and Wisniewski, 2018). The flattening of the Phillips curve has recently attracted attention by policymakers (e.g.



Notes: For this figure, we choose the demand shock to generate an output gap as measured by either the Congressional Budget Office (CBO) or the FRB's (demeaned) utilization rate. Shaded areas represent NBER recessions.

Figure 8: Partial elasticity of real marginal cost with respect to output

Brainard, 2019).

4.3.3 The welfare cost of business cycles

Building on Lucas' (1987) seminal contribution, we next discuss the implications for the welfare costs of business cycles. As Lucas famously demonstrated, the welfare benefits of eliminating business cycle fluctuations in standard models are shockingly small—and smaller than compatible with many researchers' intuition. This discrepancy has prompted work exploring the origins and the robustness of this finding. We build here on one branch of this research agenda that generates larger welfare costs of business cycles by introducing non-linearities on the supply side.[20] With such non-linearities, Jensen's inequality implies that the mean of consumption in the presence of business cycle fluctuations differs from the level of consumption that would prevail in the absence of business cycle fluctuations. In our model the mean of consumption falls permanently, thereby raising the welfare costs of business cycles.

Table 8 shows the welfare costs of business cycles in our model. We calibrate the demand shock such that the standard deviation of the log of consumption is 0.032 (as in Lucas, 1987). Relative to a calibration without business cycle fluctuations, consumption falls by 0.097 percent in the baseline calibration. At the same time, however, hours worked fall by 0.076 percent, partially offsetting

[20]The welfare costs of business cycles in Lucas (1987) arise from the curvature of the utility function. Baqaee and Farhi (2019) illustrate that the welfare costs can be very large in economies with network linkages and nonlinearities on the production side. In Barlevy (2004) decreasing returns to investment generate a permanent increase in the growth rate when business cycle fluctuations are eliminated, implying large welfare costs of business cycles. Lucas (2003) and Barlevy (2005) summarize the literature.

	$\frac{\mathbb{E}[C]-C_{ss}}{C_{ss}}$	$\frac{\mathbb{E}[n]-n_{ss}}{n_{ss}}$	Welfare costs
Model/calibration		(in percent)	
Baseline	-0.097	-0.076	0.119
Elasticity of substitution 0.5	-0.077	-0.059	0.115
22 capacity industries	-0.071	-0.055	0.107
66 capacity industries	-0.129	-0.101	0.135
Frisch labor supply elasticity 1	-0.114	-0.090	0.139
Frisch labor supply elasticity 5	-0.088	-0.068	0.107
Lucas (1987)	-	-	0.068

Notes: As in Lucas (1987), the welfare costs are measured in consumption equivalents.

Table 8: Welfare costs of business cycles

the utility loss from lower consumption. The representative household would permanently give up 0.119 percent of her steady state consumption in order to eliminate business cycles.

Alternative calibrations affect the loss in consumption and hours relative to the economy without business cycles as well as the welfare costs. In comparison to Lucas (1987) with the same coefficient of relative risk aversion, the welfare losses are typically a little less than twice as large. We note that our model does not feature endogenous capital accumulation, idiosyncratic income risk, and a number of other mechanisms, which the literature has shown to affect the welfare costs of business cycles.

5 Conclusion

In this paper we have shown that capacity constraints at the plant-level generate convex supply curves at the industry level. This convexity implies that price and production responses to demand shocks are state-dependent. Comparing industries with utilization rates below the 15th and above the 85th percentile, our non-parametric estimates indicate that, a one percent outward shift in demand raises prices by approximately 0.1 percent in the former and by 0.6 percent in the latter case. When comparing the production responses of industries below the 15th percentile to those above the 85th percentile, it approximately doubles in size. In this sense, our estimates imply that the degree of convexity is large.

Convex supply curves at the industry level matter for business cycles because the aggregate supply curve becomes convex. The aggregation necessitates additional assumptions, which affect the curvature of the aggregate supply curve. In our model, the aggregate implications of convex supply curves are of intermediate size. Depending on the calibration, the fiscal multiplier increases by 7 to 16 cents on the dollar during times of slack relative to boom times. The Phillips curve flattens during recessions and over time. For instance, the elasticity of real marginal costs with respect to output falls by 18 to 30 percent during the Volcker disinflation and by 12 to 20 percent during the Great Recession. The welfare costs of business cycles approximately double.

We add three remarks. First, our estimates imply that changes in the utilization rate *due to demand shocks* cause inflation. Without conditioning on this demand shock, however, this relationship is delicate due to simultaneity bias. One would therefore not generally expect the utilization rate to be useful for forecasting inflation, although a number of studies find that it is (Corrado and Mattey, 1997; Stock and Watson, 1999). Further, a weak unconditional relationship between capacity utilization and inflation is *not* an indication that the FRB's measures of capacity utilization are flawed.

Second, the controversy as to whether the fiscal multiplier varies with the business cycle stems in part from disagreement on which variable should be used to measure the state of the business cycle. Our findings suggest that capacity utilization is a good candidate—as does prior work by Fazzari, Morley, and Panovska (2015). At the same time, it is unlikely that capacity utilization captures all relevant notions of slack, and hence other measures should be considered as well.

Third, a limitation of this and many other papers which use data on capacity utilization is that these data are only collected for the manufacturing sector, mining, and utilities. Further, and unlike earlier work by Shapiro (1989) three decades ago, we find the FRB's current measures of capacity and utilization to be highly informative. In our assessment, it would therefore be fruitful to expand efforts to measure capacity utilization to non-manufacturing industries.

In this paper, we have shown that capacity constraints lead to economically significant nonlinearities on the supply side of the economy, and we have discussed three instances in which they matter for policy. However, convex supply curves give rise to many more policy relevant implications. They imply, for instance, that responses to *all* shocks are state dependent, and not only for government expenditures. We expect that exploring the role of capacity constraints and convex supply curves for the business cycle will likely yield many more interesting insights.

References

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83:2411–2451.

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.

Atalay, Enghin. 2017. "How Important Are Sectoral Shocks?" *American Economic Journal: Macroeconomics* 9 (4):254–280.

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

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

Baqaee, David Rezza and Emmanuel Farhi. 2019. "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem." *Econometrica* 87 (4):1155–1203.

Barlevy, Gadi. 2004. "The Cost of Business Cycles Under Endogenous Growth." *The American Economic Review* 94 (4):964–990.

———. 2005. "The cost of business cycles and the benefits of stabilization." *Economic Perspectives* 29 (1):32,49.

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

Basu, Susanto and John G Fernald. 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy* 105 (2):249–283.

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

Bigio, Saki and Jennifer La'O. 2016. "Distortions in Production Networks." Nber working papers, National Bureau of Economic Research, Inc.

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

Bils, Mark and James A. Kahn. 2000. "What Inventory Behavior Tells Us about Business Cycles." *American Economic Review* 90 (3):458–481.

Bils, Mark and Peter J. Klenow. 1998. "Using Consumer Theory to Test Competing Business Cycle Models." *Journal of Political Economy* 106 (2):233–261.

Boehm, Christoph E., Aaron Flaaen, and Nitya Pandalai-Nayar. 2019. "Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tōhoku Earthquake." *The Review of Economics and Statistics* 101 (1):60–75.

Brainard, Lael. 2019. "The Disconnect between Inflation and Employment in the New Normal." Remarks at the national tax association 49th annual spring symposium, Board of Governors of the Federal Reserve System.

Bresnahan, Timothy F. and Valerie A. Ramey. 1994. "Output Fluctuations at the Plant Level." *The Quarterly Journal of Economics* 109 (3):593–624.

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

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 1995. "Capital Utilization and Returns to Scale." In *NBER Macroeconomics Annual 1995, Volume 10*, NBER Chapters. National Bureau of Economic Research, Inc, 67–124.

Cassels, J. M. 1937. "Excess Capacity and Monopolistic Competition." *The Quarterly Journal of Economics* 51 (3):426,443.

Chamberlin, Edward. 1933. *The theory of monopolistic competition*. Harvard economic studies ... vol. XXXVIII. Cambridge: Harvard University Press.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. "When Is the Government Spending Multiplier Large?" *Journal of Political Economy* 119 (1):78–121.

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.

Corrado, Carol and Joe Mattey. 1997. “Capacity Utilization.” *Journal of Economic Perspectives* 11 (1):151–167.

Cragg, John G. and Stephen G. Donald. 1993. “Testing Identifiability and Specification in Instrumental Variable Models.” *Econometric Theory* 9 (2):222–240.

Fagnart, Jean-François, 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.

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

Hall, Robert E. 1990. “Invariance Properties of Solow’s Productivity Residual.” In *Growth/ Productivity/Unemployment: Essays to Celebrate Robert Solow’s Birthday*, edited by Peter Diamond. Cambridge, MA: MIT Press, 71–112.

———. 2009. “By How Much Does GDP Rise If the Government Buys More Output?” *Brookings Papers on Economic Activity* 40 (2 (Fall)):183–249.

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

Hulten, Charles R. 1978. “Growth Accounting with Intermediate Inputs.” *Review of Economic Studies* 45 (3):511–518.

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.

Klein, L. 1960. “Some theoretical issues in the measurement of capacity.” *Econometrica* 28 (2):272,286.

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

Levinsohn, James and Amil Petrin. 2003. “Estimating production functions using inputs to control for unobservables.” *Review of Economic Studies* 70 (243):317–341.

Lucas, Robert E. 1987. *Models of business cycles*. OX, UK: B. Blackwell.

———. 2003. “Macroeconomic Priorities.” *American Economic Review* 93 (1):1–14.

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

Morin, Norman J. and John J. Stevens. 2004. “Estimating capacity utilization from survey data.” Tech. rep.

Nakamura, Emi and Jón Steinsson. 2018. “Identification in Macroeconomics.” *Journal of Economic Perspectives* 32 (3):59–86.

Olley, G Steven and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64 (6):1263–1297.

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

Perry, George L. 1973. “Capacity in Manufacturing.” *Brookings Papers on Economic Activity* 4 (3):701–742.

Pierce, Justin R. and Emily Wisniewski. 2018. “Some Characteristics of the Decline in Manufacturing Capacity Utilization.” Feds notes, Board of Governors of the Federal Reserve System (U.S.).

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.

Robinson, Joan. 1933. *The economics of imperfect competition, by Joan Robinson.* London: Macmillan and Co., ltd.

Rotemberg, Julio J. 1982. “Sticky Prices in the United States.” *Journal of Political Economy* 90 (6):1187–1211.

Sanderson, Eleanor and Frank Windmeijer. 2016. “A weak instrument F-test in linear IV models with multiple endogenous variables.” *Journal of Econometrics* 190 (2):212–221.

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.

Smithies, Arthur. 1957. “Economic Fluctuations and Growth.” *Econometrica* 25.

Stock, James and Motohiro Yogo. 2005. *Testing for Weak Instruments in Linear IV Regression.* New York: Cambridge University Press, 80–108.

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

For online publication

A Appendix: Capacity utilization at the plant level

This Appendix discusses background information on the Quarterly Survey of Plant Capacity Utilization (QSPC) and basic facts on plants' capacity utilization using public use microdata.

Background 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. Figure A1 shows the question on the survey form. The full survey form of the Quarterly Survey of Plant Capacity Utilization is available at `https://bhs.econ.census.gov/bhs/pcu/watermark_form.pdf`. Capacity utilization rates are then obtained by dividing the market value of actual production by the estimate of full capacity production.

Why do plants produce below capacity? The survey also contains questions on why establishments produce at levels below their capacity. As Figure A1 shows, 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 2018q2, 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). These responses are summarized in Figure A2.

Item 2 **VALUE OF PRODUCTION**

A. Report market value of **actual production** for the quarter.

ACTUAL PRODUCTION . . . .

$Bil.	Mil.	Thou.

B. Estimate the market value of production of this plant as if it had been operating at **full production capability** for the quarter.

Assume:

- only machinery and equipment **in place and ready to operate**.
- normal downtime.
- labor, materials, utilities, etc. **ARE FULLY AVAILABLE**.
- the number of shifts, hours of operation and overtime pay that can be **sustained** under **normal** conditions and a **realistic** work schedule in the long run.
- the **same product mix** as the actual production.

FULL PRODUCTION CAPABILITY . . . .

$Bil.	Mil.	Thou.

C. Divide your **actual production** estimate by your **full production estimate**. Multiply this ratio by 100 to get a percentage. . . . .

Capacity Utilization

%

Is this a reasonable estimate of your utilization rate for this quarter? ☐ Yes ☐ No — *Review item 2A and 2B*

Item 3 **ACTUAL AND FULL PRODUCTION COMPARISONS**

A. FULL PRODUCTION CAPABILITY: CURRENT QUARTER VS PREVIOUS QUARTER

If your estimate of current quarter **full production capability** has changed compared to the previous quarter, mark (X) the primary reasons.

- ☐ Building capital expenditures
- ☐ Machinery capital expenditures – *Include new, replaced, or enhanced machinery*
- ☐ Building retirements
- ☐ Machinery retirements
- ☐ Price changed but product mix is the same
- ☐ Revised estimation assumption with no change in plant or operations
- ☐ Change in method of operation
- ☐ Change in product mix or product specifications
- ☐ Change in material input
- ☐ Other – *Specify*

B. ACTUAL OPERATIONS VS FULL PRODUCTION CAPABILITY

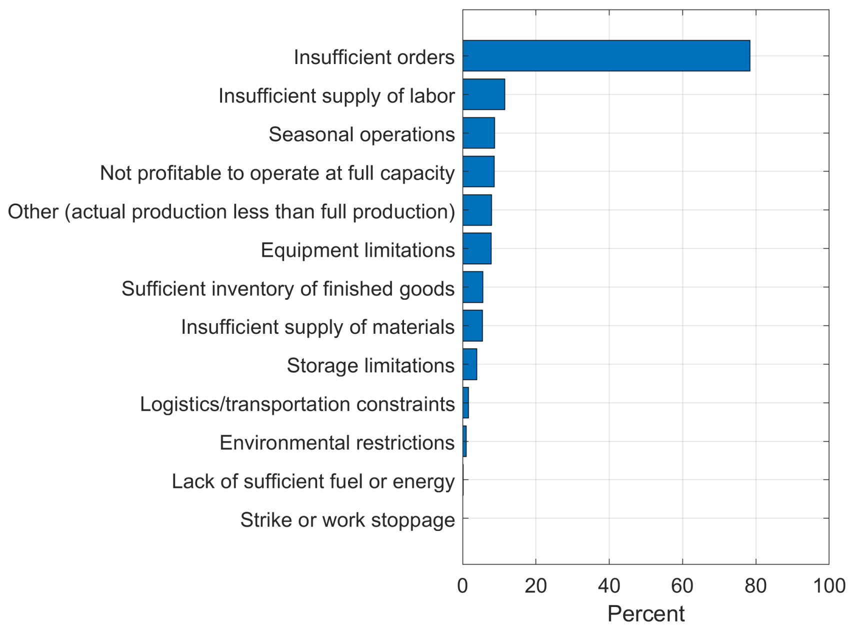
If this plant's **actual** production in the current quarter was **less** than **full production capability**, mark (X) the primary reasons.

- ☐ Not most profitable to operate at full production capability
- ☐ Insufficient supply of materials
- ☐ Insufficient orders
- ☐ Insufficient supply of local labor force/skills
- ☐ Lack of sufficient fuel or electric energy
- ☐ Equipment limitations
- ☐ Storage limitations
- ☐ Logistics/transportation constraints
- ☐ Sufficient inventory of finished goods on hand
- ☐ Strike or work stoppage
- ☐ Seasonal operations
- ☐ Environmental restrictions
- ☐ Other – *Specify*

CONTINUE WITH* Item 4 *ON PAGE 3.

31016025

Figure A1: Page 2 of survey form of the Quarterly Survey of Plant Capacity



Notes: The data are from public use data of the QSPC of the U.S. Census Bureau and are averaged from 2013q1 to 2018q2.

Figure A2: Qualitative responses

B Appendix: Proofs

Recall that we assumed in Section 2 that $\theta > 1$, that $\mathbb{E}[\omega] = 1$ and that $\mathbb{E}\left[\omega^2\right] < \infty$. We will make use of these conditions below. Further, for the proofs in this Appendix, the following limits are useful. Using L'Hôspital's rule, we have

$$\lim_{\bar{\omega}_t \to 0} \frac{\int_0^{\bar{\omega}_t} \omega dG(\omega)}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}} = \lim_{\bar{\omega}_t \to 0} \frac{\bar{\omega}_t g(\bar{\omega}_t)}{\frac{\theta-1}{\theta}(\bar{\omega}_t)^{-\frac{1}{\theta}}} = \frac{\theta}{\theta-1} \lim_{\bar{\omega}_t \to 0} \bar{\omega}_t^{1+\frac{1}{\theta}} g(\bar{\omega}_t) = 0, \tag{B1}$$

$$\lim_{\bar{\omega}_t \to \infty} (\bar{\omega}_t)^{\frac{\theta-1}{\theta}} \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} dG(\omega) = \lim_{\bar{\omega}_t \to \infty} \frac{-(\bar{\omega}_t)^{\frac{1}{\theta}} g(\bar{\omega}_t)}{-\frac{\theta-1}{\theta}(\bar{\omega}_t)^{-\frac{\theta-1}{\theta}-1}} = \frac{\theta}{\theta-1} \lim_{\bar{\omega}_t \to \infty} (\bar{\omega}_t)^2 g(\bar{\omega}_t) = 0, \tag{B2}$$

where g is the pdf of G.

Proof of Lemma 1

Lemma 1. *The utilization rate as defined in (10) 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. $u' < 0$

Proof. Using the definition of capacity and limit (B1), we can write

$$\begin{aligned} Q(q_t) &= \lim_{\bar{\omega}_t \to 0} Y(q_t, \bar{\omega}_t) \\ &= q_t \left(\lim_{\bar{\omega}_t \to 0} \frac{1}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}} \int_0^{\bar{\omega}_t} \omega dG(\omega) + \int_0^{\infty} \omega^{\frac{1}{\theta}} dG(\omega) \right)^{\frac{\theta}{\theta-1}} \\ &= q_t \Theta, \end{aligned}$$

where

$$\Theta = \left(\int_0^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega) \right)^{\frac{\theta}{\theta-1}}.$$

Then equation (10) implies that

$$u(\bar{\omega}_t) = \frac{1}{\Theta} \left(\left(\frac{1}{\bar{\omega}_t} \right)^{\frac{\theta-1}{\theta}} \int_0^{\bar{\omega}_t} \omega dG(\omega) + \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} dG(\omega) \right)^{\frac{\theta}{\theta-1}}. \tag{B3}$$

Hence, u_t is only a function of $\bar{\omega}_t$ and $u_t \geq 0$.

Regarding part 2, $\lim_{\bar{\omega}_t \to 0} u(\bar{\omega}_t) = 1$ follows directly from the definition of capacity and utilization. Further,

$$\begin{aligned} \lim_{\bar{\omega}_t \to \infty} u(\bar{\omega}_t) &= \lim_{\bar{\omega}_t \to \infty} \frac{1}{\Theta} \left(\left(\frac{1}{\bar{\omega}_t} \right)^{\frac{\theta-1}{\theta}} \int_0^{\bar{\omega}_t} \omega g(\omega) d\omega + \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} g(\omega) d\omega \right)^{\frac{\theta}{\theta-1}} \\ &= \frac{1}{\Theta} \left(\lim_{\bar{\omega}_t \to \infty} \frac{\int_0^{\bar{\omega}_t} \omega g(\omega) d\omega}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}} \right)^{\frac{\theta}{\theta-1}} = 0 \end{aligned}$$

For part 3, take the derivative of equation (B3) to obtain

$$\frac{\partial u_t}{\partial \bar{\omega}_t} = -\frac{(\bar{\omega}_t)^{\frac{1-2\theta}{\theta}}}{\Theta} \left(\left(\frac{1}{\bar{\omega}_t} \right)^{\frac{\theta-1}{\theta}} \int_0^{\bar{\omega}_t} \omega g(\omega)\, d\omega + \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} g(\omega)\, d\omega \right)^{\frac{1}{\theta-1}} \int_0^{\bar{\omega}_t} \omega g(\omega)\, d\omega, \tag{B4}$$

which is negative for $0 < \bar{\omega}_t < \infty$. It then follows that $u_t \leq 1$, completing the proof of part 1 of the lemma.

□

Proof of Proposition 1

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.*

Proof. The dynamic problem in Section 2.1.2 requires that the firms pricing decision solves

$$\max_{p_t^y(j)} \left(p_t^y(j) - mc_t\right) y_t(j)$$

subject to the constraints

$$y_t \leq q_t,$$

$$y_t(j) = \omega_t(j)\, Y_t \left[\frac{p_t^y(j)}{P_t^Y} \right]^{-\theta}.$$

Letting $\rho_t(j)$ be the multiplier on the capacity constraint, the solution requires that firms set prices according to the rule

$$p_t^y(j) = \frac{\theta}{\theta-1}\left(mc_t + \rho(j)\right), \tag{B5}$$

with $\rho(j) = 0$ if and only if $y_t = q_t$.

The threshold value $\bar{\omega}_t$ of the demand shock is

$$\bar{\omega}_t = \frac{q_t}{Y_t} \left(\frac{\theta}{\theta-1} \frac{mc_t}{P_t^Y} \right)^{\theta} \tag{B6}$$

and production of firms with $\omega_t(j) \geq \bar{\omega}_t$ is constrained by capacity while production of firms with $\omega_t(j) < \bar{\omega}_t$ is not. For constrained firms, the Lagrange multiplier is

$$\rho(j) = \frac{\theta-1}{\theta} P_t^Y \left(\frac{\omega_t(j)\, Y_t}{q_t} \right)^{\frac{1}{\theta}} - mc_t. \tag{B7}$$

Now combining the price index (3) with the price setting rule (B5) and equations (B6) and (B7) gives

$$\begin{aligned} P_t^Y &= \left(\int_0^1 \omega_t(j) \left[\frac{\theta}{\theta-1} \left(mc_t + \rho(j)\right) \right]^{1-\theta} dj \right)^{\frac{1}{1-\theta}} \\ &= \frac{\theta}{\theta-1} mc_t \left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}} \int_{\bar{\omega}_t}^{\infty} \omega_t^{\frac{1}{\theta}} dG(\omega) \right)^{\frac{1}{1-\theta}}. \end{aligned}$$

Taking logs gives the supply curve (11) with the log average markup given by

$$\tilde{\mathcal{M}}(\bar{\omega}_t) := \ln\frac{\theta}{\theta-1} - \frac{1}{\theta-1}\ln\left(\int_0^{\bar{\omega}_t}\omega_t dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)\right).$$

The log average markup as a function of $\ln u_t$ is then $\mathcal{M}(\ln u_t) := \tilde{\mathcal{M}}(\bar{\omega}(\ln u_t))$.

For part 1, note that

$$\tilde{\mathcal{M}}'(\bar{\omega}_t) = -\frac{1}{\theta}\frac{(\bar{\omega}_t)^{-\frac{1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega_t dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}. \tag{B8}$$

Further note that

$$\begin{aligned}\mathcal{M}'(\ln u_t) &= \tilde{\mathcal{M}}'(\bar{\omega}_t)\cdot\frac{\partial\bar{\omega}(u_t)}{\partial u_t}\cdot u_t \\ &= \tilde{\mathcal{M}}'(\bar{\omega}_t)\cdot\left(\frac{\partial u_t}{\partial\bar{\omega}_t}\right)^{-1}\cdot u_t\end{aligned}$$

Now plugging in equations (B8), (B4), and (B3) gives, after some algebra,

$$\mathcal{M}'(\ln u_t) = \frac{1}{\theta}\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega dG(\omega)}, \tag{B9}$$

which is greater than or equal to zero.

For part 2 note that

$$\begin{aligned}\lim_{u\to0}\mathcal{M}(\ln u_t) &= \lim_{\bar{\omega}\to\infty}\tilde{\mathcal{M}}(\bar{\omega}_t) \\ &= \ln\frac{\theta}{\theta-1} - \frac{1}{\theta-1}\ln\left(\int_0^{\infty}\omega dG(\omega) + \lim_{\bar{\omega}_t\to\infty}(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega^{\frac{1}{\theta}}dG(\omega)\right) \\ &= \ln\frac{\theta}{\theta-1} - \frac{1}{\theta-1}\ln\left(1 + \lim_{\bar{\omega}_t\to\infty}(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega^{\frac{1}{\theta}}dG(\omega)\right) \\ &= \ln\frac{\theta}{\theta-1}\end{aligned}$$

where we used the limit (B2). Further

$$\lim_{u\to1}\mathcal{M}(\ln u_t) = \lim_{\bar{\omega}\to0}\tilde{\mathcal{M}}(\bar{\omega}_t) = \ln\frac{\theta}{\theta-1} - \frac{1}{\theta-1}\ln\left(\lim_{\bar{\omega}_t\to0}(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega^{\frac{1}{\theta}}dG(\omega)\right) = \infty$$

For part 3, and using the limits (B1) and (B2), we obtain

$$\lim_{u_t\to0}\mathcal{M}'(\ln u_t) = \lim_{\bar{\omega}_t\to\infty}\frac{1}{\theta}\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega dG(\omega)} = \lim_{\bar{\omega}_t\to\infty}\frac{1}{\theta}(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega) = 0$$

and

$$\lim_{u_t\to1}\mathcal{M}'(\ln u_t) = \lim_{\bar{\omega}_t\to0}\frac{1}{\theta}\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega dG(\omega)} = \lim_{\bar{\omega}_t\to0}\frac{1}{\theta}\int_0^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}}{\int_0^{\bar{\omega}_t}\omega dG(\omega)} = \infty.$$

For part 4, take the derivative of equation (B9) to obtain

$$\mathcal{M}''(\ln u_t) = \frac{\partial\left(\frac{1}{\theta}\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega dG(\omega)}\right)}{\partial\bar{\omega}_t}\left(\frac{\partial u_t}{\partial\bar{\omega}_t}\right)^{-1}u_t.$$

Since $u_t > 0$ and $\frac{\partial u_t}{\partial \bar{\omega}_t} < 0$, the sign of $\mathcal{M}''(\ln u_t)$ is the negative of the sign of the first derivative on the right hand side. Now

$$\begin{aligned}
&\frac{\partial\left(\frac{1}{\theta}\frac{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\int_0^{\bar{\omega}_t}\omega dG(\omega)}\right)}{\partial\bar{\omega}_t} \\
= \; & \frac{\left[\frac{\theta-1}{\theta}(\bar{\omega}_t)^{-\frac{1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega) - \bar{\omega}_t g(\bar{\omega}_t)\right]\int_0^{\bar{\omega}_t}\omega dG(\omega) - \bar{\omega}_t g(\bar{\omega}_t)(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)}{\theta\left(\int_0^{\bar{\omega}_t}\omega dG(\omega)\right)^2} \\
= \; & \frac{\frac{\theta-1}{\theta}(\bar{\omega}_t)^{-\frac{1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)\int_0^{\bar{\omega}_t}\omega dG(\omega) - \bar{\omega}_t g(\bar{\omega}_t)\left[\int_0^{\bar{\omega}_t}\omega dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}}\int_{\bar{\omega}_t}^{\infty}\omega_t^{\frac{1}{\theta}}dG(\omega)\right]}{\theta\left(\int_0^{\bar{\omega}_t}\omega dG(\omega)\right)^2}.
\end{aligned} \tag{B10}$$

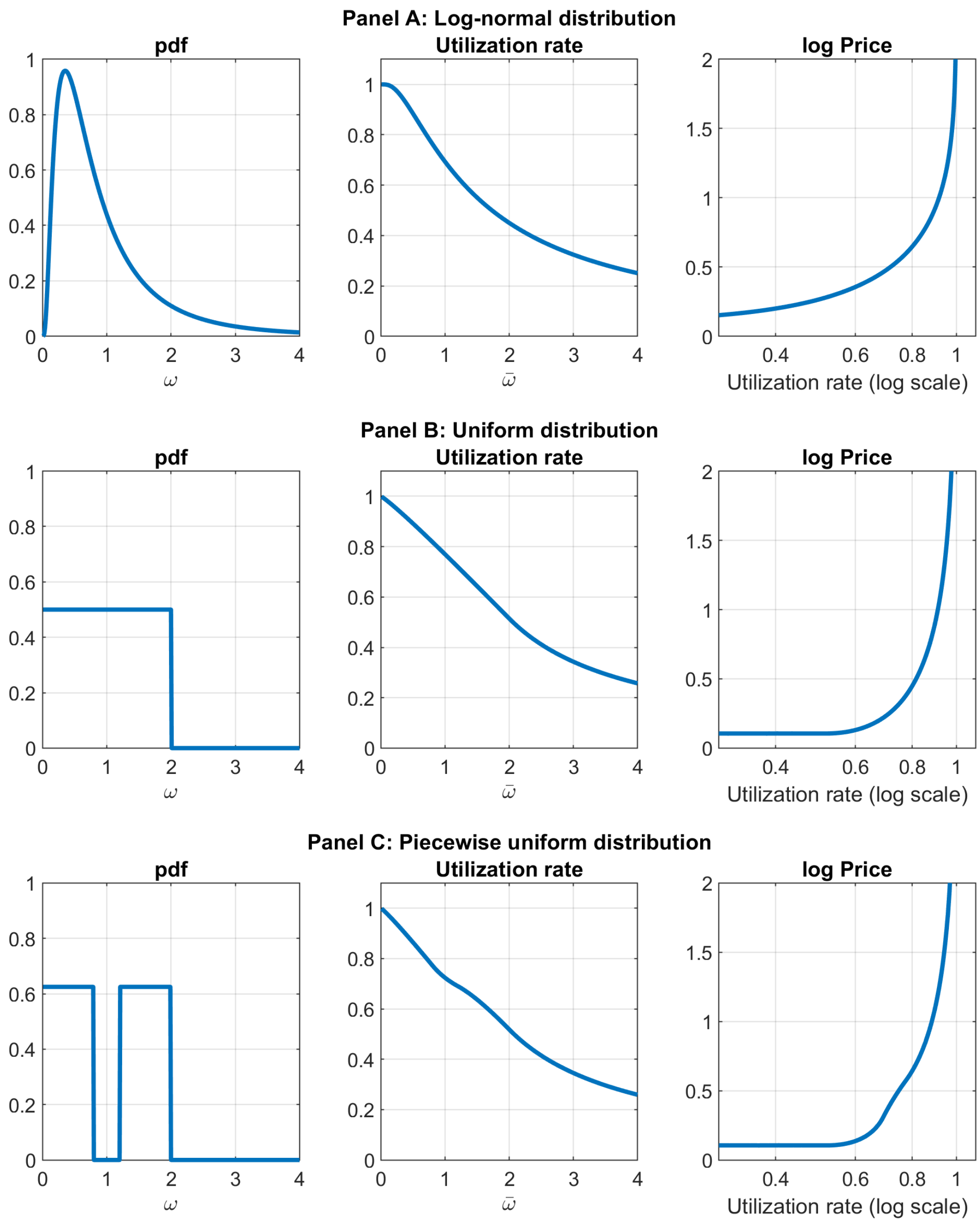
It is clear that this expression can be positive or negative, depending on the value of $g(\bar{\omega}_t)$. If $g(\bar{\omega}_t)$ is sufficiently small, the derivative on the left hand side is positive and $\mathcal{M}''$ negative. For sufficiently large $g(\bar{\omega}_t)$, the opposite is the case. Most conventional distributions result in $\mathcal{M}'' > 0$. This completes the proof.

We provide an example of a non-convex supply curve in Figure B1.

□

Example of non-convex supply curve

Figure B1 provides three additional examples of supply curves. In panel A the distribution G of demand shocks is log-normal with unit mean and variance 1. In panel B the distribution of G is uniform on the interval from 0 to 2. In Panel C the distribution of G is piecewise uniform from 0 to 0.8 and from 1.2 to 2. For panels A and B the supply curve is convex everywhere. For panel C, the supply curve is locally non-convex. This non-convexity is a result of the density of demand shocks, $g(\bar{\omega}_t)$ in equation B10, being equal to zero in the relevant range. We view this example as likely not economically relevant.



Notes: The figure provides examples of supply curve (11). The parameterizations are chosen as follows. For all figures $\theta = 10$ and marginal costs mc are set to 1. In panel A, G is log-normal with unit mean and variance 1. In panel B, G is uniform from 0 to 2. In panel C, G is piecewise uniform from 0 to 0.8 and 1.2 to 2.

Figure B1: Examples of supply curves

Proof of Proposition 2

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

$$\begin{aligned}\Delta \ln Y_{i,t} \quad = \quad & \beta_\xi\left(\ln u_{i,t-1}\right)\Delta\xi_{i,t}+\beta_\pi\left(\ln u_{i,t-1}\right)\Delta\pi_{i,t}+\beta_e\left(\ln u_{i,t-1}\right)\Delta e_{i,t}\\ & +\beta_Q\left(\ln u_{i,t-1}\right)\Delta\ln Q_{i,t}+\beta_{mc}\left(\ln u_{i,t-1}\right)\Delta\ln mc_{i,t}\\ & +\beta_{IM}\left(\ln u_{i,t-1}\right)\frac{\Delta IM_{i,t}}{Y_{i,t-1}}+\beta_{inv}\left(\ln u_{i,t-1}\right)\frac{\Delta Y_{i,t}^{inv}-\Delta Y_{i,t-1}^{inv}}{Y_{i,t-1}}+\omega_{i,t}^{Y}.\end{aligned}$$

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 $\beta_{inv}>0$. Supply curve (11) is convex if and only if $\beta_\xi'<0$ and $\beta_e'<0$. The error term is a weighted average of changes in the unobserved demand shocks $\omega_{i,t}^{C}$, $\omega_{j,i,t}^{I}$, $\omega_{j,i,t}^{M}$, and $\omega_{d,i,t}^{EX}$.

Proof. We prove a slightly more general version in which the demand elasticity is allowed to depend on the type of customer. Setting $\sigma^M=\sigma^C=\sigma^I=\sigma^F=\sigma$ gives the version in the text. In this more general version $\Delta\pi_{i,t}$ is defined as

$$\begin{aligned}\Delta\pi_{i,t} \quad = \quad & \frac{\sigma^M}{\bar\sigma}\sum_j s_{j,i,t-1}^{M}\Delta\ln P_{j,t}^{M}+s_{i,t-1}^{C}\frac{\sigma^C}{\bar\sigma}\Delta\ln P_t^C+\frac{\sigma^I}{\bar\sigma}\sum_j s_{j,i,t-1}^{I}\Delta\ln P_{j,t}^{I}\\ & +\frac{\sigma^F}{\bar\sigma}\sum_d s_{d,i,t-1}^{EX}\Delta\ln P_{d,t}^{EX,*}.\end{aligned}$$

Define the shares $s_{d,i,t-1}^{EX}=\frac{EX_{d,i,t-1}}{Y_{i,t-1}}$, $s_{j,i,t-1}^{M}=\frac{M_{j,i,t-1}}{Y_{i,t-1}}$, $s_{j,i,t-1}^{I}=\frac{I_{j,i,t-1}}{Y_{i,t-1}}$, $s_{i,t-1}^{C}=\frac{C_{i,t-1}}{Y_{i,t-1}}$, $s_{i,t-1}^{G}=\frac{G_{i,t-1}}{Y_{i,t-1}}$, $s_{i,t-1}^{EX}=\sum_d s_{d,i,t-1}^{EX}$, $s_{i,t-1}^{M}=\sum_j s_{j,i,t-1}^{M}$, and $s_{i,t-1}^{I}=\sum_j s_{j,i,t-1}^{I}$. We begin with summarizing the linearized system of equations:

Supply

$$\Delta\ln P_{i,t}^{Y}=\mathcal{M}'\left(\ln u_{i,t-1}\right)\Delta\ln u_{i,t}+\Delta\ln\left(mc_{i,t}\right)$$

Domestic final demand

$$\Delta\ln C_{i,t}=\Delta\ln\omega_{i,t}^{C}+\Delta\ln C_t-\sigma^C\Delta\ln P_{i,t}^{Y}+\sigma^C\Delta\ln P_t^C$$

$$\Delta\ln I_{j,i,t}=\Delta\ln\omega_{j,i,t}^{I}+\Delta\ln I_{j,t}-\sigma^I\Delta\ln P_{i,t}^{Y}+\sigma^I\Delta\ln P_{j,t}^{I}$$

Intermediate demand

$$\Delta\ln M_{j,i,t}=\Delta\ln\omega_{j,i,t}^{M}+\Delta\ln M_{j,t}-\sigma^M\Delta\ln P_{i,t}^{Y}+\sigma^M\Delta\ln P_{j,t}^{M}$$

Foreign demand

$$\Delta\ln EX_{d,i,t}=\Delta\ln\omega_{d,i,t}^{EX}+\Delta\ln EX_{d,t}-\sigma^F\Delta\ln P_{d,i,t}^{Y,*}+\sigma^F\Delta\ln P_{d,t}^{EX,*}$$

Utilization

$$\Delta\ln u_{i,t}=\Delta\ln Y_{i,t}-\Delta\ln Q_{i,t}$$

Price in domestic and foreign currency

$$\Delta\ln P_{i,t}^{Y}=\Delta\ln\mathcal{E}_{d,t}+\Delta\ln P_{d,i,t}^{Y,*}$$

Market Clearing

$$\begin{aligned}\Delta \ln Y_{i,t} &= \sum_j s^M_{j,i,t-1} \Delta \ln M_{j,i,t} + s^C_{i,t-1} \Delta \ln C_{i,t} + \sum_j s^I_{j,i,t-1} \Delta \ln I_{j,i,t} \\ &\quad + s^G_{i,t-1} \Delta \ln G_{i,t} + \sum_d s^{EX}_{d,i,t-1} \Delta \ln EX_{d,i,t} + \frac{\Delta Y^{inv}_{i,t} - \Delta Y^{inv}_{i,t-1}}{Y_{i,t-1}} - \frac{\Delta IM_{i,t}}{Y_{i,t-1}}\end{aligned}$$

After some algebra, this system can be solved for $\Delta \ln Y_{i,t}$ where

$$\begin{aligned}\Delta \ln Y_{i,t} &= \beta_\xi (\ln u_{i,t-1}) \Delta \xi_{i,t} + \beta_\pi (\ln u_{i,t-1}) \Delta \pi_{i,t} + \beta_e (\ln u_{i,t-1}) \Delta e_{i,t} \\ &\quad + \beta_Q (\ln u_{i,t-1}) \Delta \ln Q_{i,t} + \beta_{mc} (\ln u_{i,t-1}) \Delta \ln (mc_{i,t}) \\ &\quad + \beta_{IM} (\ln u_{i,t-1}) \frac{\Delta IM_{i,t}}{Y_{i,t-1}} + \beta_{inv} (\ln u_{i,t-1}) \frac{\Delta Y^{inv}_{i,t} - \Delta Y^{inv}_{i,t-1}}{Y_{i,t-1}} + \omega^Y_{i,t}.\end{aligned}$$

Letting

$$\bar{\sigma} = s^M_{i,t-1} \sigma^M + s^C_{i,t-1} \sigma^C + s^I_{i,t-1} \sigma^I + s^{EX}_{i,t-1} \sigma^F,$$

the coefficients satisfy

$$\begin{aligned}\beta_\xi (\ln u_{i,t-1}) &= \frac{1}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} > 0, \\ \beta_\pi (\ln u_{i,t-1}) &= \frac{\bar{\sigma}}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} > 0, \\ \beta_e (\ln u_{i,t-1}) &= \frac{\sigma^F}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} > 0, \\ \beta_Q (\ln u_{i,t-1}) &= \frac{\bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} > 0, \\ \beta_{mc} (\ln u_{i,t-1}) &= \frac{-\bar{\sigma}}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} < 0, \\ \beta_{IM} (\ln u_{i,t-1}) &= \frac{-1}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} < 0, \\ \beta_{inv} (\ln u_{i,t-1}) &= \frac{1}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} > 0.\end{aligned}$$

Further, the error term is

$$\begin{aligned}\omega^Y_{i,t} &= \frac{1}{1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1})} \cdot \\ &\quad \left(\sum_j s^M_{j,i,t-1} \Delta \ln \omega^M_{j,i,t} + s^C_{i,t-1} \Delta \ln \omega^C_{i,t} + \sum_j s^I_{j,i,t-1} \Delta \ln \omega^I_{j,i,t} + \sum_d s^{EX}_{d,i,t-1} \Delta \ln \omega^{EX}_{d,i,t} \right).\end{aligned}$$

Since

$$\begin{aligned}\beta'_\xi (\ln u_{i,t-1}) &= -\frac{\bar{\sigma} \mathcal{M}'' (\ln u_{i,t-1})}{(1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1}))^2}, \\ \beta'_e (\ln u_{i,t-1}) &= -\frac{\sigma^F \bar{\sigma} \mathcal{M}'' (\ln u_{i,t-1})}{(1 + \bar{\sigma} \mathcal{M}' (\ln u_{i,t-1}))^2},\end{aligned}$$

the supply curve is convex ($\mathcal{M}'' (\ln u_{i,t-1}) > 0$) if and only if $\beta'_\xi (\ln u_{i,t-1}) < 0$ and $\beta_e (\ln u_{i,t-1}) < 0$.

$\square$

Proof of Proposition 3

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

Proof. Since marginal costs are not directly observable, we are interested in using average unit costs as a proxy. Unit variable costs are

$$\frac{\int_0^1 p_t^v v_t(j)\,dj}{Y_t} = \frac{p_t^v}{z_t F(\kappa_t, 1)} \frac{\int_0^1 y_t(j)\,dj}{Y_t} = mc_t \frac{\int_0^1 y_t(j)\,dj}{Y_t},$$

where we used relationship $y_t(j) = z_t F(\kappa_t, 1) v_t(j)$ and equation (6). Now, using equation (B6), we obtain

$$\begin{aligned}
\int_0^1 y_t(j)\,dj &= \int_0^{\bar{\omega}_t} \omega_t Y_t \left[\frac{p_t^y(\omega)}{P_t^Y}\right]^{-\theta} dG(\omega) + \int_{\bar{\omega}_t}^{\infty} q_t dG(\omega) \\
&= Y_t \left[\frac{\frac{\theta}{\theta-1} mc_t}{P_t^Y}\right]^{-\theta} \int_0^{\bar{\omega}_t} \omega dG(\omega) + \int_{\bar{\omega}_t}^{\infty} q_t dG(\omega) \\
&= q_t \left(\frac{1}{\bar{\omega}_t} \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right).
\end{aligned}$$

Next, using equation (9), we can write

$$\frac{\int_0^1 p_t^v v_t(j)\,dj}{Y_t} = mc_t \frac{\left(\frac{1}{\bar{\omega}_t} \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)}{\left(\frac{1}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}} \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega)\right)^{\frac{\theta}{\theta-1}}}.$$

Defining $\Omega(\ln u_t) = \tilde{\Omega}(\bar{\omega}(u_t))$, where

$$\tilde{\Omega}(\bar{\omega}(u_t)) = -\ln\left(\frac{\left(\frac{1}{\bar{\omega}_t} \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)}{\left(\frac{1}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}} \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega)\right)^{\frac{\theta}{\theta-1}}}\right), \tag{B11}$$

it follows that

$$\ln mc_t = \ln \frac{\int_0^1 p_t^v v_t(j)\,dj}{Y_t} + \tilde{\Omega}(\bar{\omega}_t)$$

and hence

$$\ln P_t^Y = \mathcal{M}(\ln u_t) + \Omega(\ln u_t) + \ln UVC_{i,t},$$

where $UVC_{i,t} = \left(\int_0^1 p_t^v v_t(j)\,dj\right)/Y_t$.

We are interested in estimating $\mathcal{M}'(\ln u_t)$ and $\mathcal{M}''(\ln u_t)$, but $\ln u_t$ traces out the composite term $\Xi(\ln u_t) = \mathcal{M}(\ln u_t) + \Omega(\ln u_t)$. We will next show that $\Omega'(\ln u_t) < 0$ and $\Omega''(\ln u_t) < 0$ in the model. This implies that we estimate a lower bound for both the slope and the curvature

$$\mathcal{M}'(\ln u_i) = \Xi'(\ln u_i) - \Omega'(\ln u_i) \geq \Xi'(\ln u_i),$$

$$\frac{\mathcal{M}''(\ln u_i)}{u_i} = \frac{\Xi''(\ln u_i)}{u_i} - \frac{\Omega''(\ln u_i)}{u_i} \geq \frac{\Xi''(\ln u_i)}{u_i}.$$

Start with $\Omega(\ln u_t) = \tilde{\Omega}(\bar{\omega}(u_t))$ and differentiate both sides with respect to $\ln u_t$. This gives

$$\begin{aligned}
\Omega'(\ln u_t) &= \tilde{\Omega}'(\bar{\omega}(u_t)) \cdot \frac{\partial \bar{\omega}_t}{\partial u_t} \cdot u_t \\
&= \tilde{\Omega}'(\bar{\omega}_t) \cdot \left(\frac{\partial u_t}{\partial \bar{\omega}_t}\right)^{-1} \cdot u_t
\end{aligned} \tag{B12}$$

Now taking the derivative of equation (B11) gives

$$\tilde{\Omega}'(\bar{\omega}_t) = \frac{(\bar{\omega}_t)^{-\frac{1}{\theta}} \left[\int_{\bar{\omega}_t}^{\infty} \left((\omega_t)^{\frac{1}{\theta}} - (\bar{\omega}_t)^{\frac{1}{\theta}}\right) dG(\omega)\right] \int_0^{\bar{\omega}_t} \omega_t dG(\omega)}{\left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}} \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega)\right)\left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \bar{\omega}_t \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)},$$

which is positive.

Plugging this derivative together with equation (B4) and (B3) into equation (B12) gives

$$\begin{aligned}
\Omega'(\ln u_t) &= -\frac{(\bar{\omega}_t)^{-\frac{1}{\theta}} \left[\int_{\bar{\omega}_t}^{\infty} \left((\omega_t)^{\frac{1}{\theta}} - (\bar{\omega}_t)^{\frac{1}{\theta}}\right) dG(\omega)\right] \int_0^{\bar{\omega}_t} \omega_t dG(\omega)}{\left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^{\frac{\theta-1}{\theta}} \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega)\right)\left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \bar{\omega}_t \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)} \\
&\quad \cdot \frac{\left(\left(\frac{1}{\bar{\omega}_t}\right)^{\frac{\theta-1}{\theta}} \int_0^{\bar{\omega}_t} \omega g(\omega) d\omega + \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} g(\omega) d\omega\right)}{(\bar{\omega}_t)^{\frac{1-2\theta}{\theta}} \int_0^{\bar{\omega}_t} \omega g(\omega) d\omega} \\
&= -\frac{(\bar{\omega}_t)^{-\frac{1}{\theta}-1} \left[\int_{\bar{\omega}_t}^{\infty} \left((\omega_t)^{\frac{1}{\theta}} - (\bar{\omega}_t)^{\frac{1}{\theta}}\right) dG(\omega)\right]}{\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \bar{\omega}_t \int_{\bar{\omega}_t}^{\infty} dG(\omega)} < 0.
\end{aligned}$$

This completes the first part of the proof.

Next define the auxiliary function

$$\vartheta(\bar{\omega}_t) = -\frac{(\bar{\omega}_t)^{-\frac{1}{\theta}-1} \left[\int_{\bar{\omega}_t}^{\infty} \left((\omega_t)^{\frac{1}{\theta}} - (\bar{\omega}_t)^{\frac{1}{\theta}}\right) dG(\omega)\right]}{\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \bar{\omega}_t \int_{\bar{\omega}_t}^{\infty} dG(\omega)} \tag{B13}$$

and note that $\Omega'(\ln u_t) = \vartheta(\bar{\omega}_t(u_t))$. Then

$$\begin{aligned}
\Omega''(\ln u_t) &= \vartheta'(\bar{\omega}_t) \cdot \frac{\partial \bar{\omega}_t}{\partial u_t} \cdot u_t \\
&= \vartheta'(\bar{\omega}_t) \cdot \left(\frac{\partial u_t}{\partial \bar{\omega}_t}\right)^{-1} \cdot u_t
\end{aligned}$$

Since $u'(\bar{\omega}_t) < 0$ and $u_t > 0$, the sign of $\Omega''(\ln u_t)$ is fully determined by the sign of $\vartheta'(\bar{\omega}_t)$. Taking the derivative of equation (B13) gives

$$\begin{aligned}
\vartheta'(\bar{\omega}_t) &= \frac{\frac{1}{\theta}(\bar{\omega}_t)^{-\frac{1}{\theta}-1} \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega) \left(\bar{\omega}_t \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^2 \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)}{\left(\bar{\omega}_t \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^2 \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)^2} \\
&\quad + \frac{\left(\int_{\bar{\omega}_t}^{\infty} \left[(\omega_t)^{\frac{1}{\theta}} - (\bar{\omega}_t)^{\frac{1}{\theta}}\right] dG(\omega)\right)\left(\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + 2(\bar{\omega}_t) \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)}{(\bar{\omega}_t)^{\frac{1}{\theta}} \left(\bar{\omega}_t \int_0^{\bar{\omega}_t} \omega_t dG(\omega) + (\bar{\omega}_t)^2 \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right)^2},
\end{aligned}$$

which is greater than zero. Hence, $\Omega''(\ln u_t) < 0$. This completes the proof. $\square$

C Appendix: Model with sticky prices

In this appendix, we extend the baseline model in Section 2 by introducing sticky prices as in Rotemberg (1982). Alvarez-Lois (2004, 2006) also studies extensions of Fagnart, Licandro, and Portier (1999) to include sticky prices.

C.1 Aggregating firm

As before, a competitive aggregating firm uses a unit continuum of varieties, indexed j, as inputs into CES aggregator (1) to produce the industry's composite good. We continue to assume that these shocks are drawn independently and identically from distribution G with $\mathbb{E}[\omega] = 1$ and $\mathbb{E}\left[\omega^2\right] < \infty$.

Since the production of individual varieties is constrained by capacity, the aggregating firm cannot purchase unlimited amounts of a particular variety at a predetermined (or sticky) price. Taking prices as given, the final goods firm therefore maximizes profits subject to the production function (1) and the constraints

$$y_t(j) \leq q_t(j) \quad \text{for all } j, \tag{C1}$$

where $q_t(j)$ denotes the production capacity of variety j.

The resulting input demand curves are

$$y_t(j) = \omega_t(j)\, Y_t \left[\frac{p_t^y(j) + \rho_t(j)}{P_t^{Y,a}}\right]^{-\theta} \quad \text{for all } j, \tag{C2}$$

where $\rho_t(j)$ is the multiplier on the capacity constraint, and the industry's price index is given by

$$P_t^{Y,a} = \left(\int_0^1 \omega_t(j)\left(p_t^y(j) + \rho_t(j)\right)^{1-\theta} dj\right)^{\frac{1}{1-\theta}}. \tag{C3}$$

We refer to this price index as the *allocative* price index to distinguish it from the *measured* price index, which we will introduce below. The allocative price index is relevant for agents' decisions, but it cannot easily be constructed from the data, since the Lagrange multipliers $\rho_t(j)$ are not observed. The aggregating firm's demand is rationed whenever $\rho_t(j) > 0$.

C.2 Intermediate goods producers

The production function (4), firms' idiosyncratic production capacity (5), and marginal costs (6) are as in Section 2.1.2.

Dynamic problem As before, firms own their capital stock k, discount future profits at rate r, and maximize the present value of profits. In contrast to Section 2.1.2, however, we assume now that firms compete monopolistically and prices are sticky as in Rotemberg (1982). To keep the problem analytically tractable, we additionally assume that the demand shocks ω materialize *after* prices have been set. We allow investment to be subject to possibly non-convex adjustment costs $\phi_k(x,k)$. The firm's Bellman equation is then

$$\begin{aligned} &V\left(k, \bar{v}, z, p_{-1}^y\right) \\ &= \max_{p^y, k', \bar{v}'} \left\{\mathbb{E}_\omega\left[p^y y - p^v v - p^x x - \phi_p\left(p^y, p_{-1}^y\right) - \phi_k(x,k)\right] + \frac{1}{1+r}\mathbb{E}\left[V\left(k', \bar{v}', z', p^y\right)\right]\right\}, \end{aligned}$$

where the maximization is subject to equations (4), (5), (7), (8), and (C2) as well as the price adjustment cost function

$$\phi_p\left(p^y, p^y_{-1}\right) = \frac{\phi_p}{2}\left(\frac{p^y}{p^y_{-1}} - 1\right)^2 \Omega. \tag{C4}$$

The term Ω is a generic scaling factor that determines the units of adjustment costs (typically nominal GDP). In the problem above, we treat Ω as fixed.

Again, we assume that productivity z only has an industry-specific and an aggregate, but no firm-specific component. Since demand shocks ω materialize after prices have been set, this assumption still contains the degree of heterogeneity in the model and allows us to analytically aggregate output and prices to the industry level.

Price setting When firms set prices, they take into account that their production can be constrained by capacity. Whether a firm becomes constrained depends on how the demand shock ω materializes relative to a threshold level $\bar{\omega}$. If ω exceeds $\bar{\omega}$, firms produce at capacity, $y = q$. Otherwise they produce below capacity, $y < q$, and the demand curve (C2) holds with $\rho = 0$. A key implication of capacity constraints is that in the former case the quantity sold is locally unresponsive to price changes: Conditional on being constrained by capacity, the quantity sold is no longer decreasing in prices. This implies that the elasticity of the firms' expected output with respect to its price,

$$\tilde{\theta}\left(\bar{\omega}\right) := -\frac{\partial \ln E_\omega\left[y\right]}{\partial \ln p^y} = \theta \frac{\int_0^{\bar{\omega}} \omega dG\left(\omega\right)}{\int_0^{\bar{\omega}} \omega dG\left(\omega\right) + \bar{\omega}\int_{\bar{\omega}}^{\infty} dG\left(\omega\right)}, \tag{C5}$$

is *lower* than parameter θ in absolute terms. We call $\tilde{\theta}\left(\bar{\omega}\right)$ the *effective demand elasticity*.

The threshold value $\bar{\omega}$ takes the form

$$\bar{\omega} = \frac{q}{Y}\left[\frac{p^y}{P^{Y,a}}\right]^\theta. \tag{C6}$$

It is therefore inversely related to the industry's output Y. Holding all else equal, greater industry output Y reduces the threshold variety $\bar{\omega}$. This increases firms' probability of becoming constrained by capacity, and reduces the effective demand elasticity $\tilde{\theta}\left(\bar{\omega}\right)$. We shall see below that this mechanism is key for generating convex supply curves in this model with sticky prices.

Firms' optimal price setting requires that

$$p^y = \frac{\tilde{\theta}\left(\bar{\omega}\right)}{\tilde{\theta}\left(\bar{\omega}\right) - 1}\left(mc + \phi_p \tilde{\Psi}\left(\bar{\omega}, \ln \mathbf{a}\right)\right), \tag{C7}$$

where mc denote marginal costs as defined in equation (6). The term $\tilde{\Psi}$ is a function of current and future price changes,

$$\tilde{\Psi}\left(\bar{\omega}, \ln \mathbf{a}\right) = \frac{1}{\theta \frac{q}{\bar{\omega}} \int_0^{\bar{\omega}} \omega dG\left(\omega\right)}\left[\frac{1}{1+r}\mathbb{E}\left[\left(\frac{\left(p^y\right)'}{p^y} - 1\right)\frac{\left(p^y\right)'}{p^y}\Omega'\right] - \left(\frac{p^y}{p^y_{-1}} - 1\right)\frac{p^y}{p^y_{-1}}\Omega\right],$$

where $\ln \mathbf{a} = \left(\ln q, \ln\left(1+r\right), \ln\left(p^y\right)', \ln p^y, \ln p^y_{-1}, \ln \Omega, \ln \Omega'\right)'$ is a column vector. Note that price p^y is common for all firms and that $\tilde{\Psi}$ drops out when prices are flexible, $\phi_p = 0$.

Again, the firm's choices of k' and $\bar{v}'$ are unimportant for the estimating equation as long as we observe changes in the industry's capacity.

C.3 Industry capacity and utilization

The definition of industry capacity Q_t and the utilization rate u_t are as in Section 2.1.3. It is easy to see that Lemma 1 continues to hold, since the proof only depends on the definitions of Q_t and u_t.

Lemma C1. *The utilization rate as defined in (10) 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. $u' < 0$

Proof. See Appendix B. □

Hence, in this model with sticky prices, we can still invert the utilization rate in the threshold variety $\bar{\omega}_t$ and express industry-level aggregates in terms of the utilization rate.

C.4 The supply curve

After plugging in equations (C2) and (C6), the allocative price (C3) can be written as

$$P_t^{Y,a} = p_t^y \left(\int_0^{\bar{\omega}_t} \omega dG(\omega) + (\bar{\omega}_t)^{1-\frac{1}{\theta}} \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} dG(\omega) \right)^{\frac{1}{1-\theta}}. \tag{C8}$$

Next, use equation (C7), take logs, and invoke Lemma (C1), to obtain

$$\ln P_t^{Y,a} = \mathcal{M}_S(\ln u_t) + \varrho(\ln u_t) + \ln(mc_t + \phi_p \Psi(\ln u_t, \ln \mathbf{a}_t)), \tag{C9}$$

where

$$\begin{aligned} \tilde{\mathcal{M}}_S(\bar{\omega}_t) &= \ln \frac{\tilde{\theta}(\bar{\omega}_t)}{\tilde{\theta}(\bar{\omega}_t) - 1}, \\ \tilde{\varrho}(\bar{\omega}_t) &= \frac{1}{1-\theta} \ln \left(\int_0^{\bar{\omega}_t} \omega dG(\omega) + (\bar{\omega}_t)^{1-\frac{1}{\theta}} \int_{\bar{\omega}_t}^{\infty} \omega^{\frac{1}{\theta}} dG(\omega) \right), \end{aligned}$$

and $\mathcal{M}_S(\ln u_t) = \tilde{\mathcal{M}}_S(\bar{\omega}(u_t))$, $\varrho(\ln u_t) = \tilde{\varrho}(\bar{\omega}(u_t))$, and $\Psi(\ln u_t, \ln \mathbf{a}_t) = \tilde{\Psi}(\bar{\omega}(u_t), \ln \mathbf{a}_t)$. In the sticky price model equation (C9) is the relevant supply curve. Relative to the flexible price model in Section 2, there are three differences that we discuss next.

First, the log markup $\mathcal{M}_S(\ln u_t)$ is different from the baseline model. In this version of the model with sticky prices, the markup still depends on the utilization rate and only the utilization rate, but it arises because firms anticipate to be capacity constrained ex-post. Since the quantity demanded is inelastic conditionally on being capacity constrained, the relevant demand elasticity is $\tilde{\theta}(\bar{\omega})$, which depends on the utilization rate or, equivalently $\bar{\omega}_t$, see equation (C5). Note that we require that $\tilde{\theta}(\bar{\omega}_t) > 1$ at all times. It can be verified numerically that $\mathcal{M}_S$ is increasing and convex for conventional parameterizations, such as those used in Figure 2.

Second, the term $\varrho(\ln u_t)$ appears on the right hand side of equation (C9), which also depends on the utilization rate (and only the utilization rate). This term is absent in the flexible price version of the model. For fixed nominal prices, $\varrho(\ln u_t)$ captures rationing of the aggregating firm as mediated by the equilibrium shadow prices, see equation (C8). $\varrho(\ln u_t)$ can be interpreted as reflecting decreasing returns at the industry level. Up to a constant, $\varrho(\ln u_t)$ is identical to $\mathcal{M}(\ln u_t)$ as defined in Section 2, and hence a modified version of Proposition 1 applies.

Proposition C1. ϱ *has the following properties:*

1. $\varrho' \geq 0$

2. $\lim_{u\to 0} \varrho(\ln u) = 0$*,* $\lim_{u\to 1} \varrho(\ln u) = \infty$

3. $\lim_{u\to 0} \varrho'(\ln u) = 0$*,* $\lim_{u\to 1} \varrho'(\ln u) = \infty$

4. Without further restrictions on G, the sign of ϱ'' is generally ambiguous.

Proof. See proof of Proposition 1 in Appendix B. □

Importantly, ϱ is typically convex.

Third, the term $\Psi(\ln u_t, \ln \mathbf{a}_t)$, capturing sluggish price adjustment, appears on the right hand side of equation (C9). It depends on the general equilibrium environment and its slope and curvature are difficult to characterize without additional assumptions.

When linearizing equation (C9) around $t-1$ values, we obtain

$$\begin{aligned}\Delta \ln P_t^{Y,a} &= \left(\mathcal{M}'_S(\ln u_{t-1}) + \varrho'(\ln u_{t-1}) + \frac{\phi_p \frac{\partial \Psi}{\partial \ln u}(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}{mc_{t-1} + \phi_p \Psi(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}\right)(\Delta \ln Y_t - \Delta \ln Q_t) \\ &\quad + \frac{mc_{t-1}}{mc_{t-1} + \phi_p \Psi(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}\Delta \ln mc_t + \frac{\phi_p \left(\frac{\partial \Psi}{\partial \ln \mathbf{a}}(\ln u_{t-1}, \ln \mathbf{a}_{t-1})\right)'}{mc_{t-1} + \phi_p \Psi(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}\Delta \ln \mathbf{a}_t. \qquad \text{(C10)}\end{aligned}$$

Mirroring the previous discussion, this expression makes clear that, the slope and curvature have three distinct components, the markup $\mathcal{M}'_S$, the rationing term ϱ', and the term related to sticky prices $\frac{\partial \Psi}{\partial \ln u}(\ln u_{t-1}, \ln \mathbf{a}_{t-1})$. There are also indirect effects that arise from changes in $\Delta \ln \mathbf{a}_t$ in the last term on the right hand side, which depend on the utilization rate. Since we will argue below that the degree of price stickiness is small in our data, we will not provide a detailed discussion of these effects.

Although equation (C10) is insightful to understand the forces affecting the slope and curvature of the supply curve, it is infeasible for estimation, since the price $P_t^{Y,a}$ is not observable. It is however, the relevant object to understand the model's prediction for the evolution of quantities. In Section 3, we therefore estimate the reduced form, which does not require us to measure $\Delta \ln P_t^{Y,a}$. The industry's production response to demand shocks is then indirectly informative about the slope and curvature of this supply curve. We continue with a discussion of how to interpret estimates of the structural form.

C.5 Estimating equation

The allocative price index $P_t^{Y,a}$ is the relevant price for goods that are purchased from this industry. It is constructed, taking into account all economic costs, $\left(\int_0^1 (p_t^y(j) + \rho_t(j))\, y_t(j)\, dj\right)/Y_t$, including the shadow prices. However, $P_t^{Y,a}$ cannot be constructed from the data, because the Lagrange multipliers $\rho_t(j)$ are not observed by the statistical authorities. In the presence of rationing, it is unlikely that one can recover the relevant slope and curvature of the supply curve from the data. The details of this mis-measurement problem depend, of course, on how the price index is constructed from the underlying microdata.

We next discuss the implications of this measurement problem in the model. To do so, we construct the price index without the shadow prices and only sum the total dollar cost per unit of output, $\left(\int_0^1 p_t^y(j)\, y_t(j)\, dj\right)/Y_t$. This price index can be written as

$$P_t^{Y,m} = (p_t^y)^{1-\theta}\left(P_t^{Y,a}\right)^{\theta}\left(\int_0^{\bar{\omega}_t} \omega dG(\omega) + \bar{\omega}_t \int_{\bar{\omega}_t}^{\infty} dG(\omega)\right).$$

Note that it differs from $P_t^{Y,a}$, see equation (C8).

As discussed in Section 2.3 of the paper, a second issue is that marginal costs are not observed, and that the unit variable cost proxy differs from marginal costs,

$$UVC_t = mc_t \frac{\frac{1}{\bar{\omega}_t}\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} dG(\omega)}{\left(\frac{1}{(\bar{\omega}_t)^{\frac{\theta-1}{\theta}}}\int_0^{\bar{\omega}_t} \omega_t dG(\omega) + \int_{\bar{\omega}_t}^{\infty} (\omega_t)^{\frac{1}{\theta}} dG(\omega)\right)^{\frac{\theta}{\theta-1}}}.$$

Note that the wedge between UVC_t and mc_t is identical in this model with sticky prices and the model without sticky prices, see proof of Proposition 3 in Appendix B. After some algebra, it is then possible to write

$$\ln P_t^{Y,m} = \mathcal{M}_S(\ln u_t) + \ln\left(UVC_t + \phi_p \Psi^m(\ln u_t, \ln \mathbf{a}_t)\right),$$

where

$$\begin{aligned} \Psi^m(\ln u_t, \ln \mathbf{a}_t) &= \left(\int_0^{\bar{\omega}(u_t)} \omega dG(\omega) + (\bar{\omega}(u_t))^{1-\frac{1}{\theta}} \int_{\bar{\omega}(u_t)}^{\infty} \omega^{\frac{1}{\theta}} dG(\omega)\right)^{\frac{\theta}{1-\theta}} \\ &\quad \cdot \left(\int_0^{\bar{\omega}(u_t)} \omega dG(\omega) + \bar{\omega}(u_t) \int_{\bar{\omega}(u_t)}^{\infty} dG(\omega)\right) \tilde{\Psi}(\bar{\omega}(u_t), \ln \mathbf{a}_t). \end{aligned}$$

Now linearizing around $t-1$ values gives

$$\begin{aligned} \Delta \ln P_t^{Y,m} &= \left(\mathcal{M}_S'(\ln u_{t-1}) + \frac{\phi_p \frac{\partial \Psi^m}{\partial \ln u}(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}{UVC_{t-1} + \phi_p \Psi^m(\ln u_{t-1}, \ln \mathbf{a}_{t-1})}\right)(\Delta \ln Y_t - \Delta \ln Q_t) \qquad \text{(C11)} \\ &+ \frac{UVC_{t-1}}{UVC_{t-1} + \phi_p \Psi^m(\ln u_{t-1}, \ln \mathbf{a}_{t-1})} \Delta \ln UVC_t + \frac{\phi_p \left(\frac{\partial \Psi^m}{\partial \ln \mathbf{a}}(\ln u_{t-1}, \ln \mathbf{a}_{t-1})\right)'}{UVC_{t-1} + \phi_p \Psi^m(\ln u_{t-1}, \ln \mathbf{a}_{t-1})} \Delta \ln \mathbf{a}_t. \end{aligned}$$

Conclusions for the empirical analysis We next discuss the implications of estimating this specification under the assumptions that 1) this sticky price model as specified above is the true data generating process and the econometrician uses 2) measured prices and 3) the unit variable cost proxy instead of true marginal costs.

First, since passthrough of unit variable costs into prices is close to 1 in all our specifications, the data indicate that prices are not particularly sticky at the industry level in yearly differences. We therefore view the approximation $\phi_p \approx 0$ as reasonable. This finding is corroborated by the fact that future prices are not significant in Specification (4) of Table 5. With $\phi_p = 0$, equation (C11) becomes

$$\Delta \ln P_t^{Y,m} = \mathcal{M}_S'(\ln u_{t-1})(\Delta \ln Y_t - \Delta \ln Q_t) + \Delta \ln UVC_t.$$

Note that even if prices are flexible, $\phi_p = 0$, this model does not nest the baseline model in Section 2, because we assumed that prices are set prior to the materialization of demand shocks.

Second, relative to equation (C10), the term $\varrho'(\ln u_{t-1})$ is absent. This implies that in this model, in which prices are set prior to the materialization of demand shocks, an econometrician using 1) measured prices and 2) the unit variable cost proxy *underestimates* both the slope and the curvature of the model. This suggests that even if prices are flexible across periods, but are set before the materialization of demand shocks, we should still estimate the reduced form to check whether the implications are consistent with our estimates of the structural form.

Implications for the general equilibrium model Note also that in Section 4, we base the general equilibrium model on this version of the model, in which prices are set prior to the materialization of demand shocks. The advantage of doing so is that our estimates directly inform the markup $\mathcal{M}_S$, which we match in the calibration, see Figure G1.

D Data Appendix

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

Industrial production, capacity, and capacity utilization The series for industrial production, capacity, and capacity utilization are obtained from the Federal Reserve Board and available at `https://www.federalreserve.gov/releases/g17/ipdisk/alltables.txt`. Table D1 provides summary statistics of the utilization rate by NAICS 3-digit industry.

NBER-CES manufacturing industry database Data on prices, sales, production worker wages, material costs, energy costs, and inventories are from the NBER CES Manufacturing Industry Database. For a description of these data, see Bartelsman and Gray (1996) and Becker, Gray, and Marvakov (2016). The database is available at `http://www.nber.org/nberces/`.

BEA Input-Output Accounts Cost shares, sales shares, changes in government purchases, and changes in imports are constructed from the BEA's Input-Output accounts. Data from 1997 to today is available at Historical data is available from the BEA website here: `https://www.bea.gov/industry/input-output-accounts-data`.

BEA National Income and Product Accounts We use quantity and price indexes on personal consumption expenditures, equipment investment, and nonresidential fixed investment from the BEA's National Income and Product Accounts. These data are available at `https://www.bea.gov/national/nipaweb/DownSS2.asp`

BEA Industry Accounts Data on quantity and price indexes of downstream industries' material use from the BEA's Industry Accounts. Data from 1997 to today is available at `https://apps.bea.gov/iTable/iTable.cfm?reqid=56&step=2&isuri=1#reqid=56&step=2&isuri=1`. Historical data is available at `https://www.bea.gov/industry/historical-industry-accounts-data`. Note that since these downstream industries are not necessarily in the manufacturing sector, they are not necessarily covered by the NBER-CES manufacturing industry database.

UN Statistics Division Real GDP, the GDP deflator, both in local currency, and the nominal exchange rate are from the United Nation's Statistics Division. These data are available at `https://unstats.un.org/unsd/snaama/Downloads`.

U.S. 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`. The data are 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. The sales shares to foreign countries $s^{\mathrm{EX}}_{d,i,t}$ 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.

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

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

E Appendix: Notes on Shea's instrument

In this Appendix we provide additional notes on our version of John Shea's instrument as described in Section 3.2. In particular, we next specify a condition that guarantees that criteria (1), (2), and (3) as described in the text hold. Our starting point is equation (21).

As noted in Shea (1993b), measuring direct linkages between two industries is generally not sufficient for satisfy criteria (1), (2), and (3). Nor are ultimate cost or sales shares sufficient. Following Shea, we therefore use information from both direct and ultimate cost and sales shares. We next describe our definitions of these shares.

E.1 Demand shares

Let p_i denote the price and y_i the quantity produced by industry i. Let further $x_{j,i}$ denote industry $j's$ usage of i's output. Lastly, let d_i denote the value of final demand for the good produced by industry i.

E.1.1 Direct demand share

We define the direct demand share of industry j for industry i as

$$dds_{j,i} = \frac{p_i x_{j,i}}{\sum_j p_i x_{j,i}}.$$

While alternative definitions are sensible, we choose the denominator such that $\sum_j dds_{j,i} = 1$.

E.1.2 Ultimate demand share

Market clearing implies that

$$p_i y_i = \sum_j p_i x_{ji} + d_i = \sum_j \mu^c_{j,i} p_j y_j + d_i,$$

where $\mu^c_{j,i} = \frac{p_i x_{j,i}}{p_j y_j}$ is the cost share of i in j's output. We can then stack the system in matrix form. Using the notation

$$py = \begin{pmatrix} p_1 y_1 \\ \vdots \\ p_I y_I \end{pmatrix},\ d = \begin{pmatrix} d_1 \\ \vdots \\ d_I \end{pmatrix},\ \Gamma^c = \begin{pmatrix} \mu^c_{1,1} & \cdots & \mu^c_{1,I} \\ \vdots & \ddots & \vdots \\ \mu^c_{I,1} & \cdots & \mu^c_{I,I} \end{pmatrix},$$

we can write

$$py = d + (\Gamma^c)' py,$$

or

$$py = \left(I - (\Gamma^c)'\right)^{-1} d.$$

Based on this relationship, we define the ultimate demand share of industry j for the output of industry i as

$$uds_{j,i} = \frac{1}{p_i y_i} \cdot \left(I - \Gamma^{c\prime}\right)^{-1}_{i,j} \cdot d_j,$$

where $(I - \Gamma^{c\prime})^{-1}_{i,j}$ is the (i,j)th element of matrix $(I - \Gamma^{c\prime})^{-1}$. By construction, $\sum_j uds_{j,i} = 1$.

E.2 Cost shares

E.2.1 Direct cost share

We define industry j's direct cost share for industry i as

$$dcs_{i,j} = \frac{p_j x_{i,j}}{\sum_j p_j x_{i,j}}.$$

Notice that $\sum_j dcs_{i,j} = 1$.

E.2.2 Indirect cost share

Let va_i denote industry i's value added, then

$$p_i y_i = va_i + \sum_j p_j x_{i,j} = va_i + \sum_j \mu^s_{i,j} p_j y_j,$$

where $\mu^s_{i,j} = \frac{p_j x_{i,j}}{p_j y_j}$ is industry j's sales share to industry i. Using the notation

$$py = \begin{pmatrix} p_1 y_1 \\ \vdots \\ p_I y_I \end{pmatrix}, \; va = \begin{pmatrix} va_1 \\ \vdots \\ va_I \end{pmatrix}, \; \Gamma^s = \begin{pmatrix} \mu^s_{1,1} & \cdots & \mu^s_{I,i} \\ \vdots & \ddots & \vdots \\ \mu^s_{I,1} & \cdots & \mu^s_{I,I} \end{pmatrix},$$

we can then stack the system in matrix form and write

$$py = va + \Gamma^s py,$$

or

$$py = (I - \Gamma^s)^{-1} va.$$

The ultimate cost share of industry j for industry i is then defined as

$$ucs_{i,j} = \frac{1}{p_i y_i} \cdot (I - \Gamma^s)^{-1}_{i,j} \cdot va_j,$$

where $(I - \Gamma^s)^{-1}_{i,j}$ denotes the (i, j)th element of matrix $(I - \Gamma^s)^{-1}$. Notice that $\sum_j ucs_{i,j} = 1$.

E.3 Our version of Shea's instrument

We define our version of Shea's instrument as

$$\text{inst}^{\text{Shea}}_{i,t} = \sum_j s^M_{j,i,t-1} \mathbb{1}\left\{ \frac{\min\{\text{dds}_{j,i,t-1}, \text{uds}_{j,i,t-1}\}}{\max\{\text{dcs}_{j,i,t-1}, \text{ucs}_{j,i,t-1}, \text{dcs}_{i,j,t-1}, \text{ucs}_{i,j,t-1}\}} > 3 \right\} \Delta \ln M_{j,t}. \tag{E1}$$

Conditions (1), (2), and (3) as defined in Section 3.2 are satisfied because (1) j's the demand share from i is large relative to j's cost share from i (2) and i's cost share from j (3).

Table E1 reports the share of qualifying partner and year observations by industry (the average of the indicator in equation (E1)). As in Shea (1993a,b), for some industries no partner satisfies the criteria in any year. These are Food Manufacturing, Beverage and Tobacco Product Manufacturing, Chemical Manufacturing, Primary Metal Manufacturing, Fabricated Metal Product Manufacturing, and Transportation Equipment Manufacturing, a total of 6 out of 21.

Industry	NAICS	Share
Food Manufacturing	311	0.000
Beverage and Tobacco Product Manufacturing	312	0.000
Textile Mills	313	0.054
Textile Product Mills	314	0.054
Apparel Manufacturing	315	0.057
Leather and Allied Product Manufacturing	316	0.057
Wood Product Manufacturing	321	0.032
Paper Manufacturing	322	0.013
Printing and Related Support Activities	323	0.034
Petroleum and Coal Products Manufacturing	324	0.005
Chemical Manufacturing	325	0.000
Plastics and Rubber Products Manufacturing	326	0.021
Nonmetallic Mineral Product Manufacturing	327	0.051
Primary Metal Manufacturing	331	0.000
Fabricated Metal Product Manufacturing	332	0.000
Machinery Manufacturing	333	0.008
Computer and Electronic Product Manufacturing	334	0.002
Electrical Equipment, Appliance, and Component Manufacturing	335	0.047
Transportation Equipment Manufacturing	336	0.000
Furniture and Related Product Manufacturing	337	0.063
Miscellaneous Manufacturing	339	0.039

Table E1: Share of qualifying partner-year observations

F Appendix: Additional Results

F.1 Ad-hoc estimation of the supply curve

In this Appendix, we show the results from an *ad-hoc* estimation of the supply curve, which does not use the guidance of the model in Section 2. To do so, we estimate the specification

$$\Delta \ln P^Y_{i,t} = \beta_0 + \beta_1 \Delta \ln Y_{i,t} + \beta_2 \left(\Delta \ln Y_{i,t}\right)^2 + \text{controls} + \varepsilon_{i,t},$$

using the WID and Shea's instrument as well as their squares to address simultaneity.

The results are shown in Table F1. The key problem with using this *ad-hoc* specification is that the first stage for the squared term is uniformly weak across all three specifications ($F < 3$). Based on our instruments, it is therefore not possible to estimate the curvature of the supply curve without the structure of the model.

Dependent variable: $\Delta \ln P^Y_{i,t}$

Estimator	2SLS	2SLS	2SLS
Instrument(s):			
Main effect		WID, Shea	
Squared term		WID, Shea	
	(1)	(2)	(3)
$\Delta \ln Y_{i,t}$	0.50	0.18	0.20
	(0.37)	(0.12)	(0.12)
$(\Delta \ln Y_{i,t})^2$	-2.90	-0.88	-0.92
	(1.41)	(0.60)	(0.67)
$\Delta \ln Q_{i,t}$			-0.14
			(0.11)
$\Delta \ln \text{UVC}_{i,t}$		0.89	0.90
		(0.03)	(0.04)
R-squared	0.343	0.900	0.901
Fixed Effects	yes	yes	yes
First stage and instrument diagnostics†			
F Main effect	9.81	12.57	17.44
F Squared term	2.47	2.44	2.45
Cragg-Donald Wald F	5.61	5.87	3.94
SW F Main effect	13.27	19.80	23.77
SW F Squared term	3.30	3.32	3.29
Hansen J (p-value)	0.918	0.960	0.948

Notes: Driscoll-Kraay standard errors are reported in parentheses.
†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table F1: Ad-hoc estimation

F.2 Testing the model's coefficient restrictions

Our derived estimating equation (12) implies a number of coefficient restrictions, which we test in this appendix. The tests are based on the specification reported at the top of Table F2. The point estimates of this specification are reported in specification (6) of Table 2.

The model's restrictions are listed and tested individually in Panel A of Table F2. The model predicts that the coefficients on output and capacity sum to zero—both for the main effect and for the interaction with utilization. We cannot reject either null hypothesis. The model further predicts that the coefficient on the utilization rate is zero. Again, we cannot reject the null hypothesis. Lastly, the model predicts that the coefficient on the unit variable cost control is unity. This null hypothesis is strongly rejected. When we test these restrictions jointly in Panel B, the null hypothesis is also rejected (joint test 1). We next drop the restriction that the coefficient on unit variable costs is unity. In this case we cannot reject the null hypothesis (joint test 2).

The overall conclusion from these tests is that the model does well, except that the coefficient on the unit variable cost control is too low. However, we do not view this as a major failing of the model. A likely reason for this low coefficient is that we use unit variable cost as a proxy for marginal cost. If doing so introduces classical measurement error, the coefficient is biased towards zero.

Specification

$$\begin{aligned}\Delta \ln P_{i,t}^{Y} \;=\; & \alpha + \beta_Y \Delta \ln Y_{i,t} + \beta_{Yu} \Delta \ln Y_{i,t} \cdot (u_{i,t-1} - u_i) + \beta_u \left(u_{i,t-1} - u_i\right) \\ & + \beta_Q \Delta \ln Q_{i,t} + \beta_{Qu} \Delta \ln Q_{i,t} \cdot (u_{i,t-1} - u_i) + \beta_{\text{UVC}} \Delta \ln \text{UVC}_{i,t} + \varepsilon_{i,t}\end{aligned}$$

Panel A: Individual tests

H_0 :	p-value
$\beta_Y + \beta_Q = 0$	0.302
$\beta_{Yu} + \beta_{Qu} = 0$	0.847
$\beta_u = 0$	0.593
$\beta_{\text{UVC}} = 1$	0.000

Panel B: Joint tests

H_0 :	p-value
Joint test 1	0.000
$\beta_Y + \beta_Q = 0$	
$\beta_{Yu} + \beta_{Qu} = 0$	
$\beta_u = 0$	
$\beta_{\text{UVC}} = 1$	
Joint test 2	0.350
$\beta_Y + \beta_Q = 0$	
$\beta_{Yu} + \beta_{Qu} = 0$	
$\beta_u = 0$	

Note: These tests are Wald tests and based on Driscoll-Kraay standard errors.

Table F2: Testing the model's coefficient restrictions

F.3 Heterogeneity

In this appendix we explore cross-industry heterogeneity in the slope and curvature of the supply curve. We do so by assigning each industry to one of two groups and then estimate equation (13) while allowing the two groups of industries to have different coefficients for all right hand side variables except the fixed effects. The instruments are the World Import Demand and Shea's instrument as well as interactions of both with $u_{i,t-1} - \bar{u}_i$. The coefficients on each instrument differ by industry group in the first stage.

Specification (1) in Table F3 shows a split by durability, where durability is defined as in the North American Industry Classification System (NAICS), see Table D1. The differences between durable and nondurable goods producing industries are small and not statistically significant (p-values not reported). In specification (2) we split industries by average utilization rate. Again, there are no statistically significant differences between these two groups, although the curvature estimate is slightly higher in the high average utilization group. Note that supply curves slope up and are convex in all groupings.

F.4 Estimates of the reduced form

Linear model Specifications (1) and (2) of Table F4 show estimates of the linear reduced form 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 potentially 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 around 0.9.

Robustness We next consider a number of robustness checks. In specifications (1) of Table F5, 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 -2.55), 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 coefficient on $\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ is -2.56 and that on $\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$ is -17.53. Hence, that the production response depends on the initial utilization rate is robust to including or dropping a large number controls.

F.5 Consistency of estimates of structural and reduced form

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 F1 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 the estimated supply elasticity. We plot this latter response

Dependent variable: $\Delta \ln P^Y_{i,t}$

	(1) By durability		(2) By average utilization rate	
	nondurable	durable	low	high
$\Delta \ln Y_{i,t}$	0.16	0.20	0.26	0.21
	(0.07)	(0.06)	(0.07)	(0.10)
$\Delta \ln Y_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	1.15	0.82	1.04	1.29
	(0.39)	(0.37)	(0.31)	(0.45)
$u_{i,t-1} - \bar{u}_i$	0.05	-0.04	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.06)
$\Delta \ln Q_{i,t}$	-0.12	-0.11	-0.26	-0.13
	(0.09)	(0.09)	(0.10)	(0.16)
$\Delta \ln Q_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-0.35	-1.26	-1.19	-1.48
	(0.77)	(0.35)	(0.43)	(0.61)
$\Delta \ln \mathrm{UVC}_{i,t}$	0.86	0.94	0.83	0.92
	(0.03)	(0.03)	(0.03)	(0.03)
R-squared	0.907		0.905	
Fixed Effects	yes		yes	
First stage and instrument diagnostics†				
F Main effect	27.03	116.29	68.42	32.10
F Interaction	54.36	29.94	47.45	10.09
Cragg-Donald Wald F	7.681		4.456	
SW F Main effect	63.68	55.87	34.60	16.05
SW F Interaction	85.22	24.46	54.23	12.33
Hansen J (p-value)	0.714		0.849	

Notes: The 2SLS estimates are based on equation (13), but allow the two groups of industries to have different coefficients for all right hand side variables except the fixed effects. The instruments are the World Import Demand and Shea's instrument as well as interactions of both with $u_{i,t-1} - \bar{u}_i$ and an indicator for the category (nondurable/durable, low/high average utilization). Durability is defined at the 3-digit NAICS level. Industries in the low/high utilization groups are separated at the mean. Driscoll-Kraay standard errors are reported in parentheses. Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s^{\mathrm{EX}}_{d,i,t-1}$).

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table F3: Heterogeneity

for three alternative demand elasticities, $\sigma = 1$, 2, and 3. As is clear from the figure, both estimates are broadly consistent with one another if one believes that the demand elasticity is in this range. On the other hand, and as discussed in Appendix C, measurement problems can imply that the estimates of structural and reduced form deliver different conclusions even if the true data generating process is as specified by the model. Our preferred interpretation is therefore that the estimates of structural and reduced form are qualitatively consistent with one another, without taking a stand on whether they are quantitatively consistent.

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

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	0.81	0.77	0.60	0.87	0.85
	(0.08)	(0.07)	(0.11)	(0.42)	(0.13)	(0.13)
$\Delta\pi_{i,t}$	-0.03	-0.03	-0.10	-0.07	-0.12	-0.12
	(0.11)	(0.11)	(0.12)	(0.13)	(0.12)	(0.12)
$\Delta e_{i,t}$	1.94	1.83	0.94	1.17	0.82	0.84
	(0.60)	(0.48)	(1.05)	(1.23)	(1.01)	(1.01)
$\Delta \ln Q_{i,t}$	0.69	0.72	0.64	0.67	0.63	0.63
	(0.07)	(0.06)	(0.07)	(0.10)	(0.07)	(0.07)
$\Delta \ln \mathrm{UVC}_{i,t}$	-0.08	-0.07	-0.10	-0.10	-0.10	-0.10
	(0.05)	(0.06)	(0.04)	(0.05)	(0.04)	(0.04)
$\left(\Delta Y_{i,t}^{\text{inv}} - \Delta Y_{i,t-1}^{\text{inv}}\right)/Y_{i,t-1}$		1.00	0.89	0.92	0.87	0.87
		(0.14)	(0.15)	(0.18)	(0.15)	(0.15)
$\Delta IM_{i,t}/Y_{i,t-1}$		-0.10	0.04	0.12	0.00	0.01
		(0.08)	(0.09)	(0.19)	(0.09)	(0.09)
R-squared	0.691	0.735	0.826	0.824	0.826	0.826
Fixed Effects	no	no	yes	yes	yes	yes
			First stage and instrument diagnostics			
F Main effect				7.23	288.86	159.92
Hansen J (p-value)						0.567

Notes: The estimates are based on equation (18). Driscoll-Kraay standard errors are reported in parentheses. Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share ($\sum_d s_{d,i,t-1}^{\mathrm{EX}}$).

Table F4: Estimates of the linear reduced form

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

Estimator	2SLS					
Instruments						
Main effect	WID, Shea					
Interaction $(\cdot (u_{i,t-1} - \bar{u}_i))$	WID, Shea					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\xi_{i,t}$	0.79	0.92	0.81	0.77	0.91	0.84
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)
$\Delta\xi_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-3.60	-3.54	-3.55	-3.36	-2.66	-2.44
	(0.88)	(0.99)	(0.98)	(1.07)	(1.06)	(1.23)
$\Delta e_{i,t}$	0.38	0.56	0.27	0.69	0.86	0.86
	(0.86)	(1.28)	(0.90)	(0.81)	(0.82)	(0.79)
$\Delta e_{i,t} \cdot (u_{i,t-1} - \bar{u}_i)$	-20.45	-22.00	-20.26	-18.04	-18.66	-15.92
	(5.50)	(6.73)	(5.43)	(5.32)	(4.60)	(4.48)
$(u_{i,t-1} - \bar{u}_i)$	-0.21	0.00	-0.21	-0.24	-0.17	-0.21
	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)	(0.05)
$\Delta \ln P^Y_{i,t+1}$			0.05			0.05
			(0.03)			(0.02)
$\Delta \ln Y_{i,t-1}$				0.12		0.12
				(0.02)		(0.03)
R-squared	0.842	0.798	0.844	0.848	0.863	0.868
Fixed Effects	yes	yes	yes	yes	yes	yes
Drop $\Delta \ln \text{UVC}_{i,t}$ and interaction	yes	no	no	no	no	no
Drop $\Delta \ln Q_{i,t}$ and interaction	no	yes	no	no	no	no
All second order terms	no	no	no	no	yes	yes
Other controls	yes	yes	yes	yes	yes	yes
First stage and instrument diagnostics†						
F Main effect	108.56	143.37	109.90	104.32	72.47	69.43
F Interaction	24.30	17.43	18.23	18.17	16.75	17.38
Cragg-Donald Wald F	38.74	37.26	38.38	35.52	36.18	33.78
SW F Main effect	148.15	203.16	149.71	143.28	99.82	96.34
SW F Interaction	33.77	26.03	25.24	25.93	24.36	25.84
Hansen J (p-value)	0.652	0.388	0.543	0.828	0.990	0.989

Notes: The estimates are based on equation (18). Fixed effects include industry fixed effects, time fixed effects, and time fixed effects interacted with industries' lagged foreign sales share $(\sum_d s^{\text{EX}}_{d,i,t-1})$. Other controls include $\Delta\pi_{i,t}$, $\Delta \ln Q_{i,t}$, $\Delta \ln \text{UVC}_{i,t}$, $\Delta IM_{i,t}/Y_{i,t-1}$, $\left(\Delta Y^{\text{inv}}_{i,t} - \Delta Y^{\text{inv}}_{i,t-1}\right)/Y_{i,t-1}$, and their interactions with $u_{i,t-1} - \bar{u}_i$. Driscoll-Kraay standard errors are reported in parentheses.

†: F is the standard F-statistic. For details on the Cragg-Donald statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

Table F5: Robustness of the reduced form

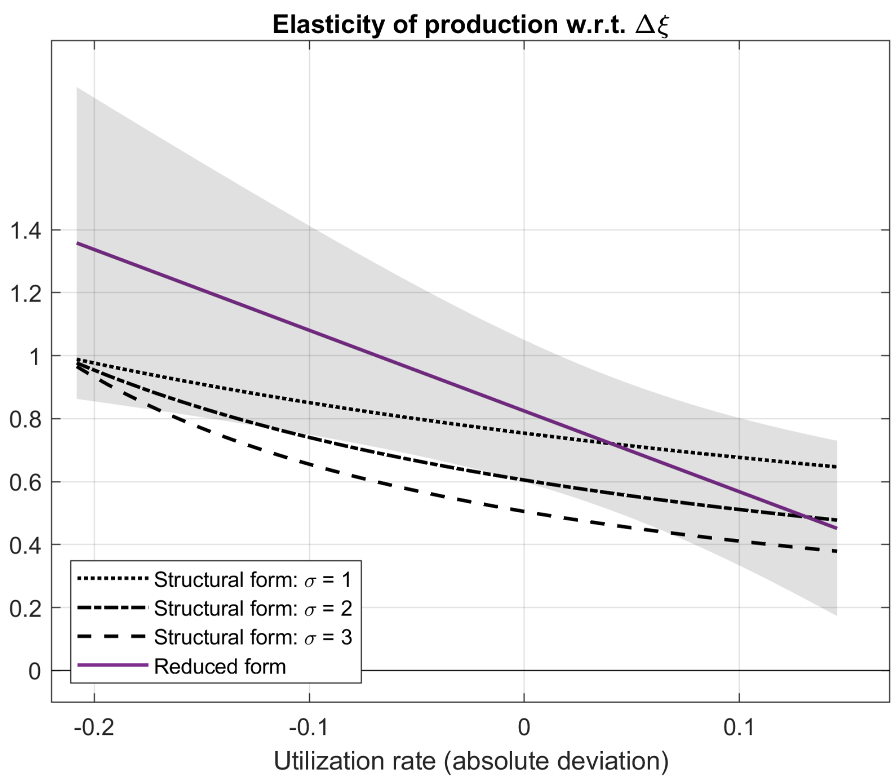


Figure F1: The production response to demand shocks

G General equilibrium model

G.1 Model setup

G.1.1 Households

Letting σ denote the intertemporal elasticity of substitution, and η the Frisch labor supply elasticity, households maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{(C_t)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \chi \frac{\sum_{i=1}^{I} (n_{i,t})^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} \right]$$

subject to the budget constraint

$$P_t^Y C_t = b_t \sum_i w_{i,t} n_{i,t} + \Pi_t - \tau_t^l.$$

In our model the aggregate shock b_t is the source of business cycle fluctuations. Since this shock scales the marginal utility of consumption, it can be interpreted as a demand shock. Relative to a shock that enters the utility function, b_t does not impact welfare calculations through a direct effect. Of course, it is related to the labor wedge (see e.g. Karabarbounis, 2014). In the budget constraint, P_t^Y is the price of output, Π_t are profits and τ_t^l lump-sum taxes. Sector i's wage is denoted by $w_{i,t}$ and labor by $n_{i,t}$. Households could trade an uncontingent bond to price the interest rate, but we omit this part for simplicity. Optimal behavior implies that

$$\phi (n_{i,t})^{\frac{1}{\eta}} = b_t (C_t)^{-\frac{1}{\sigma}} \frac{w_{i,t}}{P_t^Y}. \tag{G1}$$

G.1.2 Final aggregating firm

The final aggregating firm produces output Y_t, which is sold to the final consumer and to the government.It uses a constant elasticity of substitution (CES) aggregator with elasticity ϑ to assemble the final good from intermediates,

$$Y_t = \left(\int X_t(\ell)^{\frac{\vartheta-1}{\vartheta}} d\ell \right)^{\frac{\vartheta}{\vartheta-1}}.$$

Demand for output from intermediary ℓ is

$$X_t(\ell) = Y_t \left(\frac{P_t^X(\ell)}{P_t^Y} \right)^{-\vartheta},$$

and the price index

$$P_t^Y = \left(\int \left(P_t^X(\ell)\right)^{1-\vartheta} d\ell \right)^{\frac{1}{1-\vartheta}}.$$

We next introduce a unit mass of intermediaries, indexed ℓ, which buy goods from all industries, differentiate them, and sell them to the final aggregating firm. In our model, their only role is to add a markup over marginal costs. In a more general model, these intermediaries could have sticky prices. We will need these intermediaries to discuss the mapping from output to real marginal costs (see Section 4.3.2), in which we will not impose that these intermediates set a constant markup.

Each intermediary ℓ uses the a CES aggregator

$$X_t(\ell) = \left(\sum_i \psi_i^{\frac{1}{\varepsilon}} (X_{i,t}(\ell))^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where the sum is over all the different industries. Cost minimization implies that

$$X_{i,t}(\ell) = \psi_i X_t(\ell) \left(\frac{P_{i,t}^{Y,a}}{mc_t(\ell)} \right)^{-\varepsilon}$$

where $mc_t(\ell)$ denotes the marginal cost of intermediary ℓ,

$$mc_t(\ell) = \left(\sum_i \psi_{i,t} \left(P_{i,t}^{Y,a} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}.$$

In the baseline model prices are flexible, and each intermediary ℓ charges a constant markup over marginal costs,

$$P_t^X(\ell) = \frac{\vartheta}{\vartheta - 1} mc_t(\ell).$$

Note next that all intermediaries have the same technology. It follows that $mc_t(\ell) = mc_t$ and hence $P_t^X(\ell) = P_t^X = P_t^Y$. Then

$$\frac{mc_t}{P_t^Y} = \left(\sum_i \psi_i \left(\frac{P_{i,t}^{Y,a}}{P_t^Y} \right)^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}} \tag{G2}$$

and

$$X_{i,t} = \psi_i Y_t \left(\frac{P_{i,t}^{Y,a}}{P_t^Y} \frac{P_t^Y}{mc_t} \right)^{-\varepsilon}. \tag{G3}$$

Further, in the baseline model $\frac{mc_t}{P_t^Y} = \frac{\vartheta - 1}{\vartheta}$.

G.1.3 Capacity industries

The capacity industries in this general equilibrium extension are a simplified version of those modelled in Appendix C. The main difference, is that capacity $q_{i,t}$ and $\kappa_{i,t}$ are treated as fixed here, while they are endogenously chosen Appendix C. Each capacity industry is populated by a unit continuum of monopolistic competitors. An aggregating firm maximizes

$$P_{i,t}^Y Y_{i,t} - \int_0^1 p_{i,t}^y(j)\, y_{i,t}(j)\, dj$$

subject to the production function

$$Y_{i,t} = \left(\int_0^1 \omega_{i,t}(j)^{\frac{1}{\theta}} \left(y_{i,t}(j) \right)^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}}$$

and the capacity constraint

$$y_{i,t}(j) \le q_{i,t}(j).$$

Optimal behavior implies that

$$y_{i,t}(j) = \omega_{i,t}(j)\, Y_t \left[\frac{p_{i,t}^y(j) + \rho_{i,t}(j)}{P_{i,t}^{Y,a}} \right]^{-\theta},$$

and

$$P_{i,t}^{Y,a} = \left(\int_0^1 \omega_{i,t}\left(j\right) \left(p_{i,t}^{y}\left(j\right) + \rho_{i,t}\left(j\right)\right)^{1-\theta} dj \right)^{\frac{1}{1-\theta}},$$

where $\rho_{i,t}\left(j\right)$ is the multiplier on the capacity constraint.

Once monopolistic competitors are aggregated to the industry level (see Section 2), we obtain the following set of equations at the industry level. First, price setting implies that

$$\frac{P_{i,t}^{Y,a}}{P_t^Y} = \frac{\tilde{\theta}_{i,t}}{\tilde{\theta}_{i,t} - 1} \frac{1}{\frac{p_{i,t}^{y}}{P_{i,t}^{Y,a}}} \frac{1}{z_{i,t}\kappa_{i,t-1}^{\alpha}} \frac{p_{i,t}^{v}}{P_t^Y} \tag{G4}$$

where the effective demand elasticity is

$$\tilde{\theta}_{i,t} = \theta \frac{\int_0^{\bar{\omega}_{i,t}} \omega_{i,t} dG\left(\omega_{i,t}\right)}{\int_0^{\bar{\omega}_{i,t}} \omega_{i,t} dG\left(\omega_{i,t}\right) + \bar{\omega}_{i,t} \int_{\bar{\omega}_{i,t}}^{\infty} dG\left(\omega_{i,t}\right)} \tag{G5}$$

and the relative price is

$$\frac{p_{i,t}^{y}}{P_{i,t}^{Y,a}} = \left(\int_0^{\bar{\omega}_{i,t}} \omega_{i,t} dG\left(\omega_{i,t}\right) + \bar{\omega}_{i,t}^{1-\frac{1}{\theta}} \int_{\bar{\omega}_{i,t}}^{\infty} \omega_{i,t}^{\frac{1}{\theta}} dG\left(\omega_{i,t}\right) \right)^{\frac{1}{\theta-1}}. \tag{G6}$$

Second, the threshold variety $\bar{\omega}_{i,t}$ is determined by the condition

$$Y_{i,t} = \frac{q_{i,t}}{\bar{\omega}_{i,t}} \left[\frac{p_{i,t}^{y}}{P_{i,t}^{Y,a}} \right]^{\theta} \tag{G7}$$

Third, demand for the variable input bundle $v_{i,t}$ is

$$z_{i,t}\left(\kappa_i\right)^{\alpha} v_{i,t} = E_{\omega}\left[y_{i,t}\right] \tag{G8}$$

where

$$E_{\omega}\left[y_{i,t}\right] = \frac{q_{i,t}}{\bar{\omega}_{i,t}} \int_0^{\bar{\omega}_{s,t}} \omega_{i,t} dG\left(\omega_{i,t}\right) + q_{i,t} \int_{\bar{\omega}_{i,t}}^{\infty} dG\left(\omega_{i,t}\right) \tag{G9}$$

G.1.4 Neoclassical (non-capacity) industries

Monopolistic competitors have the linear production function

$$y_{i,t} = z_{i,t} v_{i,t}, \tag{G10}$$

where $v_{i,t}$ is a bundle of variable inputs as discussed above. Non-capacity industries charge the markup $\frac{\theta_{nc}}{\theta_{nc}-1}$ so that

$$\frac{p_{i,t}^{Y}}{P_t^Y} = \frac{\theta_{nc}}{\theta_{nc} - 1} \frac{1}{z_{i,t}} \frac{p_{i,t}^{v}}{P_t^Y}. \tag{G11}$$

G.1.5 The variable input bundle

Both capacity and non-capacity industries produce output from a variable input bundle that is composed of labor and materials

$$v_{i,t} = \left[\omega_{n,i}^{v} \left(n_{i,t}\right)^{\frac{\lambda-1}{\lambda}} + \omega_{M,i}^{v} \left(M_{i,t}\right)^{\frac{\lambda-1}{\lambda}} \right]^{\frac{\lambda}{\lambda-1}}.$$

The materials bundle $M_{i,t}$ is itself a CES aggregate of the form

$$M_{i,t} = \left(\sum_j \left(\omega_{i,j}^M\right)^{\frac{1}{\mu}} \left(m_{i,j,t}\right)^{\frac{\mu-1}{\mu}} \right)^{\frac{\mu}{\mu-1}}.$$

We use the notation that $m_{i,j,t}$ denotes material shipments from industry j to industry i. The parameters λ and μ are elasticities of substitution and ω_n^v, ω_M^v, and $\omega_{i,j}^M$ weights. The demand curves for inputs are

$$n_{i,t} = \omega_{n,i}^v v_{i,t} \left(\frac{w_{i,t}}{p_{i,t}^v}\right)^{-\lambda}, \tag{G12}$$

$$M_{i,t} = \omega_{M,i}^v v_{i,t} \left(\frac{P_{i,t}^M}{p_{i,t}^v}\right)^{-\lambda}, \tag{G13}$$

$$m_{j,i,t} = \omega_{j,i}^M M_{i,t} \left(\frac{P_{j,t}^{Y,a}}{P_{i,t}^M}\right)^{-\mu}, \tag{G14}$$

and the price indexes are

$$p_{i,t}^v = \left(\omega_n^v \left(w_{i,t}\right)^{1-\lambda} + \omega_M^v \left(P_{i,t}^M\right)^{1-\lambda}\right)^{\frac{1}{1-\lambda}}, \tag{G15}$$

$$P_{i,t}^M = \left(\sum_j \omega_{j,i}^M \left(P_{j,t}^{Y,a}\right)^{1-\mu} \right)^{\frac{1}{1-\mu}}. \tag{G16}$$

G.1.6 Market clearing and government

The market for aggregate output clears,

$$Y_t = C_t + G_t, \tag{G17}$$

where G_t is government expenditures on output. These expenditures are financed with a contemporaneous lump-sum tax, $G_t = \tau_t^l$. Further, industries' output is sold for final use $(X_{i,t})$ and as intermediates downstream,

$$Y_{i,t} = X_{i,t} + \sum_j m_{j,i,t}. \tag{G18}$$

G.2 Equilibrium

Let I^c and I^{nc} denote the number of capacity and non-capacity industries, respectively. The total number of industries is $I = I^c + I^{nc}$.

An equilibrium in this economy is a set of prices

$$\left\{\frac{P_{i,t}^{Y,a}}{P_t^Y}\right\}_{i=1}^I, \left\{\frac{p_{i,t}^v}{P_t^Y}\right\}_{i=1}^I, \left\{\frac{w_{i,t}}{P_t^Y}\right\}_{i=1}^I, \left\{\frac{P_{i,t}^M}{P_t^Y}\right\}_{i=1}^I, \left\{\frac{p_{i,t}^y}{P_{i,t}^{Y,a}}\right\}_{i=1}^{I^c},$$

quantities

$$Y_t, c_t, \{Y_{i,t}\}_{i=1}^I, \{v_{i,t}\}_{i=1}^I, \{n_{i,t}\}_{i=1}^I, \{M_{i,t}\}_{i=1}^I, \{E_\omega\left[y_{i,t}\right]\}_{i=1}^{I^c}, \{X_{i,t}\}_{i=1}^I, \{m_{j,i,t}\}_{i=1,j=1}^{I,J},$$

threshold varieties $\{\bar{\omega}_{i,t}\}_{i=1}^{I^c}$, effective demand elasticities $\left\{\tilde{\theta}_{i,t}\right\}_{i=1}^{I^c}$ such that households and firms optimize and markets clear. This requires that equations (G1), (G2), (G3), (G4), (G5), (G6), (G7), (G8), (G9), (G10), (G11), (G12), (G13), (G14), (G15), (G16), (G17), and (G18) hold and that $\frac{mc_t}{P_t^Y} = \frac{\vartheta - 1}{\vartheta}$.

G.3 Calibration

We calibrate our model in two steps. First, we set a number of parameters to (external) conventional values in the literature. Second, we choose a other parameters to match the estimated industry-level supply curve in the data. We also discuss below which of the 71 industries in the data we calibrate as capacity and which we calibrate as non-capacity industries.

Table G1 summarizes the benchmark calibration. We highlight the role of two parameters. First, we choose a Frisch labor supply elasticity of 2. This value is greater than most micro estimates, but in line with the argument in Hall (2009). Second, we assume that the substitutability of inputs is very limited in the short run and set the elasticities of substitution to 0.05. This choice is based on recent estimates by Atalay (2017), Boehm, Flaaen, and Pandalai-Nayar (2019), and others who document very limited substitutability between inputs at business cycle frequencies.

Parameter	Description	Value	Notes
σ	Intertemporal elasticity of substitution	0.75	standard
η	Frisch labor supply elasticity	2	Hall (2009)
θ_{nc}	Parameterizes markup in nc-industries	10	standard
θ	Parameterizes markup in c-industries	14.84	internally calibrated, see text
ϑ	Parameterizes markup in final goods sector	10	standard
s_Y^G	Share of government expenditure in GDP	0.2	National Accounts
$\varepsilon, \mu, \lambda$	Elasticities of substitution	0.05	see text
$\{z_i, \kappa_i\}$	Productivity parameters	1	normalization
V	Variance of distribution G	1.75	internally calibrated, see text
$\{k_i^c\}$	Capital stock in c-industries		internally calibrated, see text
$\{\psi_i\}$	Weight in final output bundle		match expenditure share in industry i
$\{\omega_{n,i}^v\}$	Labor weight in variable input aggregator		match industry i's labor cost share
$\{\omega_{j,i}^M\}$	Weights in materials aggregator		match industry j's material cost share from i

Table G1: Baseline calibration

Weights in various CES aggregators are calibrated to expenditure and cost shares, which we take from the 2007 Use Tables of the Bureau of Economic Analysis. In particular, we use (1) the final expenditure shares in output to calibrate $\{\psi_i\}$. We use (2) the share of labor in total variable costs to calibrate $\{\omega_{n,i}^v\}$, and then set $\omega_{M,i}^v = 1 - \omega_{n,i}^v$. Lastly, we (3) calibrate $\{\omega_{j,i}^M\}$ such that j's material cost share from i equals the measured share in the data.

In a second step, we choose the elasticity θ and the variance of the log-normal distribution G of idiosyncratic demand shocks ω for monopolistic competitors in capacity industries such that the supply curve matches the estimate in the data. The mean of distribution G is normalized to 1. When doing so, we also choose industries' capital stocks k_i^c to center their utilization rate (equivalently $\bar{\omega}_{i,t}$) at the portion of the supply curve that has the curvature as estimated in the data. As Figure G1 shows, the fit between model and data is almost exact.

We next discuss which industries we calibrate as capacity industries. Since the Survey of Plant Capacity is limited to manufacturing industries, we have to decide whether our estimates apply to industries not covered by the survey. The Survey of Plant Capacity aims at measuring a capital-based notion of

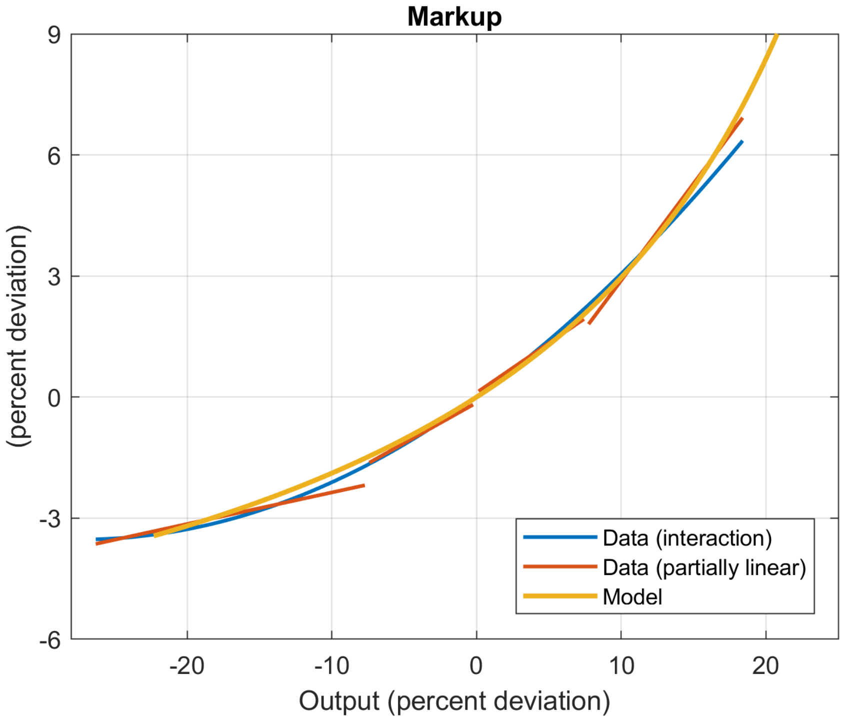


Figure G1: Calibration of the industry level supply curve

capacity utilization (see Section 3.1.1 and Appendix A). We then classify an industry as a capacity industries if this notion of capital-based capacity utilization appears to exist in the industry. Table G2 shows our classification.

Table G2: Calibration of industry types

IO code	Industry name	Calibration: Sample	Calibration: Baseline	Calibration: All private	Analogy to capital-based capacity utilization
111CA	Farms	0	1	1	Major land use, U.S. Department of Agriculture, World Bank
113FF	Forestry, fishing, and related activities	0	1	1	Major land use, U.S. Department of Agriculture, World Bank
211	Oil and gas extraction	1	1	1	in sample
212	Mining, except oil and gas	1	1	1	in sample
213	Support activities for mining	1	1	1	in sample
22	Utilities	0	1	1	Capacity Utilization: Electric and gas utilities, FRB
23	Construction	0	1	1	Capital-based capacity utilization
321	Wood products	1	1	1	in sample
327	Nonmetallic mineral products	1	1	1	in sample
331	Primary metals	1	1	1	in sample
332	Fabricated metal products	1	1	1	in sample
333	Machinery	1	1	1	in sample
334	Computer and electronic products	1	1	1	in sample
335	Electrical equipment, appliances, and components	1	1	1	in sample
3361MV	Motor vehicles, bodies and trailers, and parts	1	1	1	in sample
3364OT	Other transportation equipment	1	1	1	in sample
337	Furniture and related products	1	1	1	in sample
339	Miscellaneous manufacturing	1	1	1	in sample
311FT	Food and beverage and tobacco products	1	1	1	in sample
313TT	Textile mills and textile product mills	1	1	1	in sample
315AL	Apparel and leather and allied products	1	1	1	in sample
322	Paper products	1	1	1	in sample
323	Printing and related support activities	1	1	1	in sample
324	Petroleum and coal products	1	1	1	in sample
325	Chemical products	1	1	1	in sample
326	Plastics and rubber products	1	1	1	in sample
42	Wholesale trade	0	0	1	
441	Motor vehicle and parts dealers	0	0	1	
445	Food and beverage stores	0	0	1	
452	General merchandise stores	0	0	1	
4A0	Other retail	0	0	1	
481	Air transportation	0	1	1	Transportation capacity
482	Rail transportation	0	1	1	Transportation capacity
483	Water transportation	0	1	1	Transportation capacity
484	Truck transportation	0	1	1	Freight capacity, JOC Truckload capacity index, JOC.com
485	Transit and ground passenger transportation	0	1	1	Transportation capacity
486	Pipeline transportation	0	1	1	Transportation capacity
487OS	Other transportation and support activities	0	1	1	Transportation capacity
493	Warehousing and storage	0	1	1	Transportation capacity
511	Publishing industries, except internet (includes software)	0	0	1	
512	Motion picture and sound recording industries	0	0	1	
513	Broadcasting and telecommunications	0	0	1	
514	Data processing, internet publishing, and other information services	0	0	1	
521CI	Federal Reserve banks, credit intermediation, and related activities	0	0	1	
523	Securities, commodity contracts, and investments	0	0	1	
524	Insurance carriers and related activities	0	0	1	
525	Funds, trusts, and other financial vehicles	0	0	1	
HS	Housing	0	1	1	Capital-based capacity utilization
ORE	Other real estate	0	1	1	Capital-based capacity utilization
532RL	Rental and leasing services and lessors of intangible assets	0	1	1	Capital-based capacity utilization
5411	Legal services	0	0	1	
5415	Computer systems design and related services	0	0	1	
5412OP	Miscellaneous professional, scientific, and technical services	0	0	1	

Continued on next page

Table G2 – continued from previous page

IO code	Industry name	Calibration			Analogy to capital-based capacity utilization
		Sample	Baseline	All private	
55	Management of companies and enterprises	0	0	1	
561	Administrative and support services	0	0	1	
562	Waste management and remediation services	0	1	1	Capital-based capacity utilization
61	Educational services	0	0	1	
621	Ambulatory health care services	0	1	1	Capital-based capacity utilization
622	Hospitals	0	1	1	Occupancy rates
623	Nursing and residential care facilities	0	1	1	Occupancy rates
624	Social assistance	0	0	1	
711AS	Performing arts, spectator sports, museums, and related activities	0	1	1	Occupancy rates
713	Amusements, gambling, and recreation industries	0	1	1	Occupancy rates
721	Accommodation	0	1	1	Occupancy rates
722	Food services and drinking places	0	1	1	Occupancy rates
81	Other services, except government	0	0	1	
GFGD	Federal general government (defense)	0	0	0	
GFGN	Federal general government (nondefense)	0	0	0	
GFE	Federal government enterprises	0	0	0	
GSLG	State and local general government	0	0	0	
GSLE	State and local government enterprises	0	0	0	

G.4 Mapping from output to real marginal costs

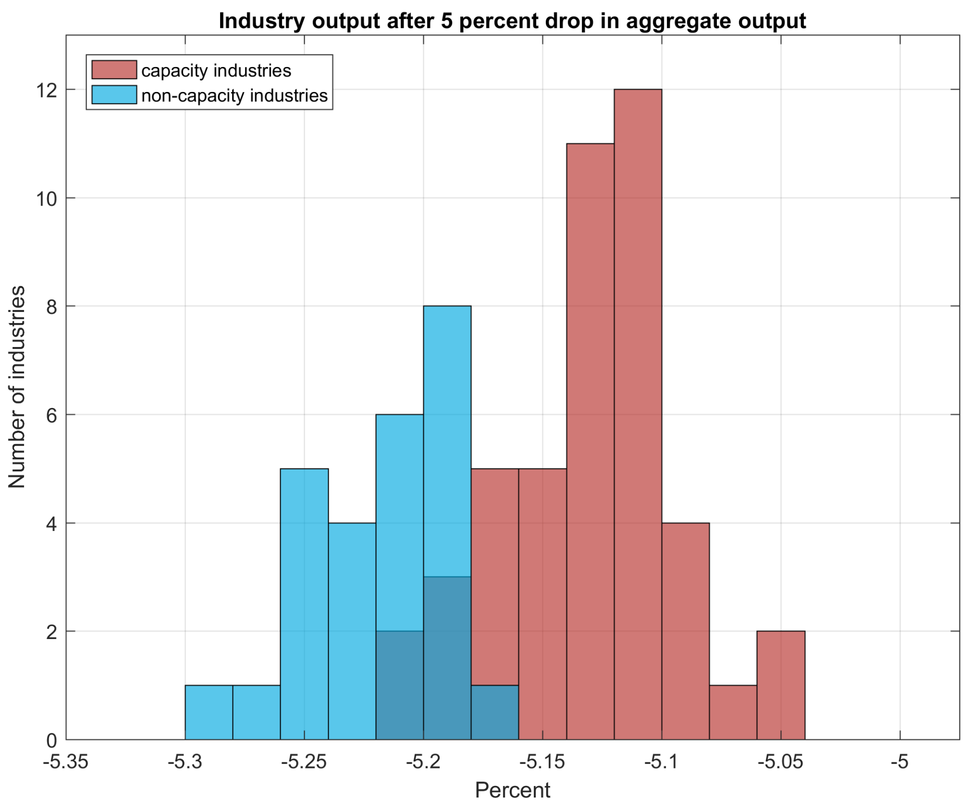
Real marginal cost as a function of aggregate output $\frac{mc_t}{P_t^Y}(Y_t)$ is defined by equations (G1), (G2), (G3), (G4), (G5), (G6), (G7), (G8), (G9), (G10), (G11), (G12), (G13), (G14), (G15), (G16), and (G18), while holding the marginal utility of consumption (or consumption C_t) constant.

The partial elasticity is then defined as in Section 4.3.2. In special cases, this mapping can be calculated analytically. For the full model, we compute it numerically.

G.5 Additional results

Heterogeneity in industries' responses to aggregate shocks As discussed in Section 4.2, industries' differential responses to aggregate shocks affect the curvature of the aggregate supply curve in the presence of non-linearities. The reason is that curvature has more bite when an industry's response to a shock is large. We demonstrate here that this channel is quantitatively unimportant in our model and calibration.

Figure G2 shows a histogram of industries' responses to a shock, which reduces aggregate output by 5 percent. As the figure shows, all industries' output falls by slightly more than 5 percent, thereby limiting this heterogeneity channel in shaping the aggregate supply curve. Capacity industries' output falls by slightly less. The reason is as follows. Since the supply curves of capacity industries are upward-sloping, a drop in demand reduces their price. Further, non-capacity industries operate constant returns to scale technologies, so their prices are not directly impacted by the fall in output. Substitution towards capacity industries explains why their output falls slightly less.



Notes: This figure is based on the benchmark calibration as described in Appendix G.3

Figure G2: Heterogeneity