Are supply curves convex? Implications for state-dependent responses to shocks*

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

[1]UT Austin

This draft: October 2017

First draft: April 2017

PRELIMINARY

Abstract

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

JEL Codes: F14, F16, F23

Keywords: State-Dependence, Fiscal Policy, Monetary Policy, Exchange Rate Passthrough, Slack, Capacity Utilization, Plant Capacity

*`chris.e.boehm@gmail.com` and `npnayar@utexas.edu`. We would like to thank Mark Aguiar, Saroj Bhattarai, Javier Cravino, Oleg Itskhoki, Andrei Levchenko, Hikaru Saijo, Chris Sims, and Jaume Ventura as well as seminar participants at Princeton and Houston for valuable comments and suggestions. 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, the Board of Governors, or its research staff.

1 Introduction

Is monetary and fiscal policy more or less effective in recessions? A recent literature in macroeconomics suggests that policy effectiveness depends on whether the economy is in recession or expansion at the time of the shock. For instance, **?** argue that the fiscal multiplier may be as high as 2.5 during recessions, but is around zero or even negative during expansions. A number of studies have put forward similar evidence, whereas others have challenged their results. As of now, no consensus has emerged as to whether such state dependence exists and, if so, whether it is quantitatively important.

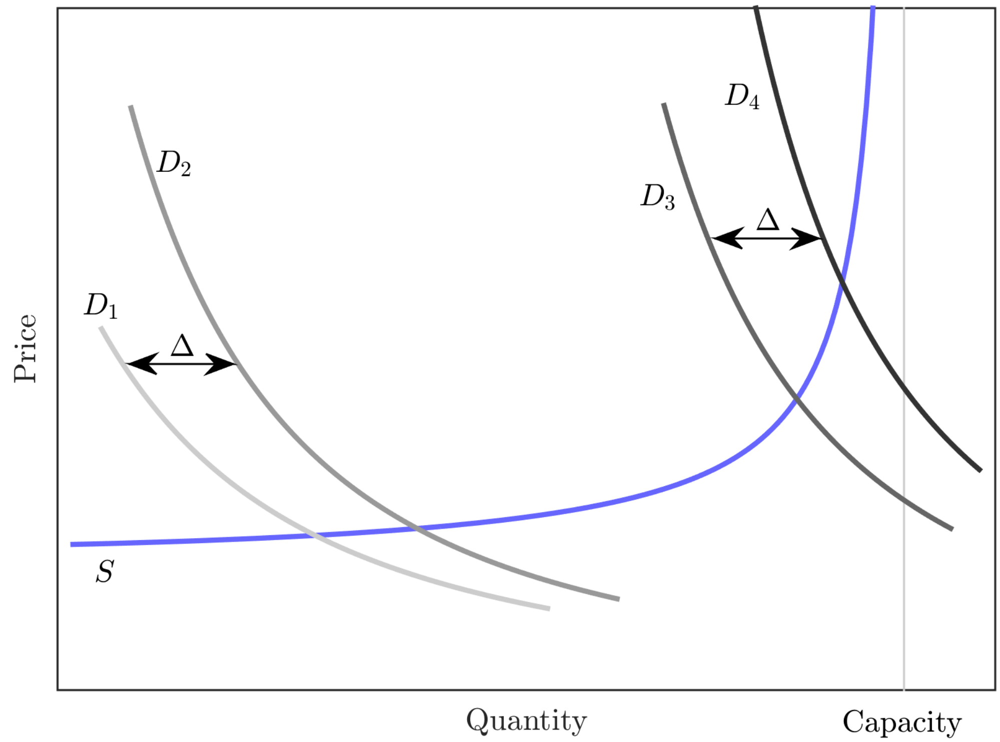
In light of its relevance for policymaking this paper revisits this question. Unlike earlier work, we study one prominent mechanism that can generate state dependent responses to shocks: the convexity of supply curves. At low levels of factor utilization supply curves are flat. As a result greater demand predominantly raises output and has small effects on prices. Stabilization policies that shift out demand are effective in such environments. Conversely, if factors of production are highly utilized, additional demand mostly raises prices with little effect on output. Figure 1 illustrates this mechanism.

We find strong evidence for convex supply curves. Our empirical analysis shows that production responds much more when the initial level of capacity utilization is low. Further, industries with high initial capacity utilization experience price increases when subjected to positive demand shocks while industries with low utilization rates do not. These results hold for two different demand shocks, one based on appropriately purified exchange rate movements and another based on variation in defense spending.

Using confidential Census data, we begin our analysis with a plausibility check of whether capacity is a likely determinant of firms' responsiveness to shocks. We show that a significant fraction of plants in a given year produces at their self-reported level of full capacity. These plants are presumably slow at adjusting production when subjected to a positive demand shock. Further, a large share of plants produce below capacity, at times far below. We interpret this as evidence of a putty-clay type scale choice that requires plants to choose their maximum production capacity far in advance. Unsurprisingly, plant capacity utilization is procyclical.

Equipped with these basic facts, we develop a model to guide our empirical analysis. Building on earlier work by Fagnart, Licandro, and Portier (1999) and consistent with the facts on plant

Figure 1: State dependent responses to shocks



capacity utilization, the main assumption is that firms invest into a set of factors that are fixed in the short run. Once chosen, these fixed factors determine maximum productive capacity and when demand for these firms materializes sufficiently high, they become capacity constrained and their production unresponsive to shocks. Our framework permits simple aggregation to the industry level, where it generates a convex supply curve.

Imposing this additional structure has a number of advantages. First, it allows us to be precise about the relevant measure of the state of the economy that determines subsequent responsiveness to shocks. In our framework the capacity utilization rate—the level of actual production divided by full capacity production—is a sufficient statistic for this state. Since the Federal Reserve estimates and publishes such capacity utilization rates we do not need to rely on ad-hoc measures that may or may not affect responses to shocks.

Second, the model yields estimating equations which allow us to directly assess whether the data supports the hypothesis that supply curves are convex. We conduct our analysis at the industry-level which has the advantage of providing more variation and more precise estimates than previous work based on aggregate data.

Finally, the structural approach implies that our findings are relevant for all types of demand shocks, although we use alternatively appropriately purified exchange rate variation or defense spending for as our identifying variation. The convexity of supply curves implies that responses to fiscal shocks, monetary shocks, and other policy interventions that have a demand shock component will also exhibit state dependence. That said, the cost of moving to the industry level obviously implies that we cannot directly test whether the U.S. economy as a whole responds differentially to shocks over the course of the business cycle.

Our empirical results support the view that responses to shocks are state dependent and that supply curves are convex. We conduct our analysis at the 3-digit NAICS level, using capacity utilization measures from the Federal Reserve Board. Further, we develop a method of purifying industry-specific effective exchange rate shocks from certain confounding factors potentially posing a threat to identification. Based on our preferred estimates, production increases by 4.69 percent after a one percent depreciation of the effective exchange rate when the initial capacity utilization rate is at the 10th percentile of the distribution. In contrast, at the 90th percentile of utilization, this elasticity is significantly lower with a value of merely 1.11. Further, and consistent with the convexity of supply curves, price responses are significantly larger when the initial utilization rate is high.

To ensure that our results are robust we extend our analysis to fiscal shocks. We construct defense spending at the industry level from the military prime contract files similar to Boehm (2016). Further, we address endogeneity concerns by using a Bartik-type instrument in the spirit of Nakamura and Steinsson (2014). Though noisy, the estimated impulse response functions confirm our earlier results. Production expands more at low initial levels of capacity utilization and prices rise only when the capacity utilization is initially high. Overall our results provide a remarkably clear picture on the state dependence of responses to demand shocks at the industry level.

Our paper complements earlier work on this topic. Auerbach and Gorodnichenko (2012, 2013a,b) provide evidence for large fiscal multipliers in times of economic slack whereas Ramey and Zubairy (2014) argue that the multiplier is below unity regardless of the state of the business cycle. Similar disagreement exists for monetary policy where Tenreyro and Thwaites (2016) find that policy is less effective in recessions while Santoro et al. (2014) provide evidence for the opposite. Further work on state-dependent responses to shocks includes Weise (1999), Peersman and Smets (2005), Lo and Piger (2005), Baum, Poplawski-Ribeiro, and Weber (2012), Owyang, Ramey, and Zubairy (2013), Fazzari, Morley, and Panovska (2015) and Jordà, Schularick, and Taylor (2017). Unlike these studies, we conduct our empirical analysis at the industry level to obtain more statistical power.[1]

A large literature in macroeconomics has studied the role of capital and capacity utilization for the transmission of business cycle shocks. These include Greenwood, Hercowitz, and Huffman (1988), Bils and Cho (1994), Cooley, Hansen, and Prescott (1995), Burnside and Eichenbaum (1996), Fagnart, Licandro, and Sneessens (1997), Gilchrist and Williams (2000), and others. Shapiro (1989) studies the Federal Reserve Board's series of capacity utilization as predictors of inflation, as do Stock and Watson (1999). On the theory side our work is closest to Michaillat (2014) who develops the idea that slack in the labor market (or a convex labor supply curve) leads to state-dependent fiscal multipliers.

We begin in section 2 with establishing basic facts on capacity utilization at the plant level. In section 3 we develop a simple illustrative model that features such underutilization of capacity and use it to motivate our empirical strategy. After discussing the data we present our baseline empirical results in section 4, using exchange rate variation for identification. We extend the empirical analysis to fiscal shocks in section 5. Section 6 concludes.

2 Capacity utilization at the plant level

We begin with presenting three basic facts on plants' capacity utilization using microdata from the Quarterly Survey of Plant Capacity Utilization (QSPC). The survey is conducted by the U.S. Census Bureau and funded jointly by the Federal Reserve Board and the Department of Defense. The sample is drawn from all U.S. manufacturing and publishing plants with 5 or more production employees. Among other things, establishments are asked about the market value of their *actual production* and the estimated market value of their *full production capacity*. Respondents are asked to construct this estimate under the following assumptions: 1) only the current functional

[1] Our work is also related to a separate literature on time varying volatility versus responsiveness. See Bachmann and Moscarini (2011), Vavra (2014), and Berger and Vavra (2017).

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.[2] Capacity utilization rates are then constructed by dividing the market value of actual production by the estimate of full capacity production.

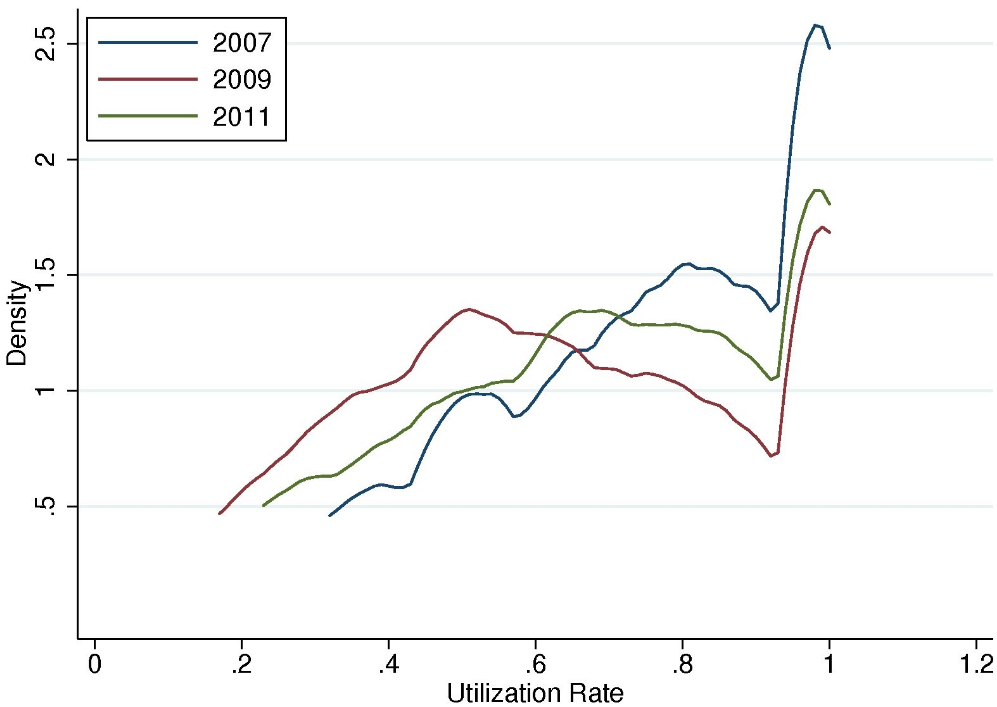
Figure 2 plots the density estimates of utilization rates for the years 2007, 2009 and 2011. Utilization rates display substantial cross-sectional variation and three facts emerge from the figure. First, a significant fraction of establishments produces at full capacity. In all three years, a discrete mass of establishments bunches at a utilization rate of unity. Second, an even larger share of plants produces below their reported capacity, and frequently far below with utilization rates around 0.2. Finally, capacity utilization at the plant level is highly procyclical. In 2007 a large fraction of plants produced at full capacity and the density displays a mode at around 0.8. By 2009 the entire distribution of utilization rates has shifted to the left and has a modal point at approximately 0.5. The 2011 density reflects partial recovery relative to 2009 but utilization rates are still below those of 2007.

Neither of these three facts is surprising. We report them because they pose difficulties to standard theories according to which firms typically operate at their optimum scale.[3] In contrast, the data appear to support a putty-clay-type view according to which firms first choose their production capacity and subsequently produce either at or below full capacity. Although our analysis thus far does not directly speak to this, it is conceivable that plants with utilization rates of unity are constrained by their existing production capacity and would produce more if they could. In sections 4 and 5 we will test directly whether capacity constraints restrict responses to demand shocks, but due to data limitations we will do so at the industry rather than the plant level. In the next section, we use the empirical facts on capacity utilization to motivate the choice of a putty-clay type production function.

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

[3] Note that our data reports utilization rates on plants and not firms.

Figure 2: Densities of plant capacity utilization



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

3 Model

The main objective of this theoretical analysis is to develop a theory that (1) can generate a convex supply curve and (2) will guide our empirical analysis. We begin with a simple illustrative model and subsequently consider generalizations with greater realism.

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

3.1 Aggregating firm

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

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

The weights v in the aggregator are firm-specific and time-varying. In this baseline version we assume for simplicity that they are drawn independently and identically from distribution G with finite third moment and unit mean.

Taking prices as given, the final goods firm maximizes profits subject to the production function (1). This problem yields the factor demand curves

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

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

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

3.2 Intermediate goods producers

Motivated by our analysis in section 2, we depart from standard theory by assuming that a firm's capacity can limit production in the short run. Following Fagnart, Licandro, and Portier (1999), we assume that the firm has to decide ex-ante on the maximum of variable factors, b_t, that it can employ (or process) in the short run. Since variable factors include primarily production workers and intermediate inputs, b_t has a natural interpretation as workstations or processing capacity of intermediates. To maintain clear notation, we drop the index j throughout this section.

We assume that *firm's idiosyncratic production capacity* takes the form

$$q_t = z_t F\left(k_t, b_t\right) \tag{4}$$

where z_t is productivity, k_t is capital, and F is constant returns to scale. The firm's actual production x_t is

$$x_t = q_t \frac{l_t}{b_t} = z_t F(\kappa_t, 1)\, l_t \tag{5}$$

where $\kappa_t = \frac{k_t}{b_t}$ and l_t represents the firm's choice of a bundle of factors that are variable in the short run. Production x_t is limited in the short run because the variable factors l_t cannot exceed b_t. Note that the marginal product of l_t, $z_t F(\kappa_t, 1)$, is constant in the short run and determined by z_t and κ_t. Letting w_t denote the price of input bundle l_t, short run marginal costs are

$$mc_t = \frac{w_t}{z_t F(\kappa_t, 1)}. \tag{6}$$

Firms own their capital stock k and maximize the present value of profits. Investment is subject to (possibly nonconvex) adjustment costs $\phi(i, k)$. The firm's Bellman equation is then

$$V(k, b, z, v) = \max_{x,i,l,b'} \left\{ px - wl - p_i i - \phi(i, k) + \frac{1}{1+r} \mathbb{E}\left[V\left(k', b', z', v'\right) \right] \right\}$$

where the maximization is subject to

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

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

and equations (2) to (5). Equation (7) is the capacity constraint and (8) is the standard capital accumulation equation. In this simple version of the model, we assume that productivity z only has an aggregate (i.e. economy wide) and an industry-specific component, but no firm-specific component.

If the firm operates below its capacity limit, it sets prices at a constant markup over marginal costs. Once it hits its capacity limit, however, it begins to raise the markup so as to equate the quantity demanded to its production capacity. Formally,

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

where ψ is the multiplier on the capacity constraint (7). In this baseline version of the model rising markups are the key channel that generates the convex supply curve. Below we discuss a number of alternative mechanisms that lead to such convexity. These include rationing (in the presence of sticky prices) and kinks in the cost function (for instance due to a shift premium). Notice that only the firm's production and price setting decision are relevant for our purposes, so we do not discuss here the firm's investment decision or its choice of b'.

Since the only source of heterogeneity is the idiosyncratic demand shock v, there exist a threshold variety $\bar{v}$ above which a firm hits the capacity constraint. A lower value of $\bar{v}$ thus indicates that more firms are capacity constrained. It turns out that $\bar{v}$ plays a critical role for characterizing the degree to which the industry uses its productive capacity. For instance, it is possible to write the industry's output as

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

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

3.3 Capacity and utilization

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

$$Q_t := \lim_{\bar{v}_t \to 0} X\left(q_t, \bar{v}_t\right).$$

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

$$u_t := \frac{X\left(q_t, \bar{v}_t\right)}{Q_t}. \tag{9}$$

This definition has several attractive properties. First, it is constructed very similar to its empirical counterpart. As we will discuss below, the Federal Reserve constructs its utilization measures by dividing an index of industrial production by its estimate of capacity. Further, the utilization rate u_t can be expressed as an (appropriately constructed) average of firms' idiosyncratic utilization rates $u_t\left(j\right)$ which are constructed, as in the Quarterly Survey of Plant Capacity Utilization, by dividing the market value of actual production by the market value of full capacity production $u_t\left(j\right) := \frac{x(j)}{q_t} = \frac{p_t x_t(j)}{p_t q_t}$.

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

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

The lemma highlights that the industry's utilization rate is only a function of the threshold value $\bar{v}_t$ 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 $\bar{v}_t = \bar{v}(u_t)$. We will make extensive use of this property, both for the remainder of the theoretical analysis and when taking the model to the data.

3.4 The supply curve

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

$$\ln P_t = \Omega(u_t) + \ln(mc_t), \tag{10}$$

where Ω is only a function of the industry's utilization rate. A key insight here is that the supply curve is not a function of output alone, but of utilization, or equivalently of output relative to capacity. If equation (10) is the true model, empirical approaches that disregard the dependence on capacity will suffer from omitted variable bias. In our empirical application below we will show that this bias appears quantitatively significant.

Proposition 1. Ω *has the following properties:*

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

Because Ω is increasing in utilization everywhere, the industry's supply curve (10) is upward-sloping. 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 Ω and its derivative tend to infinity. Conversely, when the utilization rate tends to zero, fewer and fewer suppliers are capacity constrained. As a result Ω tends to $\ln \frac{\theta}{\theta-1}$ and its derivative to zero. While $\Omega(u)$ is convex everywhere for many choices of G, it is possible to construct examples in which it is locally concave. Thus, whether Ω is convex in the relevant range of utilization remains an empirical question which we will address below.

We briefly turn to a measurement problem that affects estimation and inference. Identifying the slope and curvature of Ω typically requires controlling for marginal costs, which are not observed. Instead, a commonly taken route is to control for unit costs which are observed. Unfortunately, in our framework marginal costs and unit costs differ. Further, the wedge between them is a function of utilization which can lead to biased estimates of Ω' and Ω''.

More specifically, and letting $w_t L_t$ the industries' total expenditure on bundle l_t, marginal costs can be expressed as $\ln mc_t = \ln \frac{w_t L_t}{X_t} + \Psi(u_t)$, where Ψ is only a function of u. Hence, the industry's supply curve (10) becomes

$$\ln P_t = \Omega(u_t) + \Psi(u_t) + \ln \frac{w_t L_t}{X_t}. \tag{11}$$

Thus, variation in u does not identify Ω', but $\Omega' + \Psi'$ and a similar argument applies to Ω''. The structural approach we take allows us to sign these biases.

Proposition 2. *$\Psi \leq 0$, $\Psi' \leq 0$, $\Psi'' \leq 0$.*

Hence, empirical strategies that aim to estimate Ω' and Ω'' based on equation (11), exhibit a downward bias (for both the slope and the curvature of Ω). It is thus possible that supply curves are upward-sloping and convex, but the researcher finds no evidence supporting this, not even in large samples. Naturally, this raises the bar of for measuring state dependent responses to shocks in the data.

3.5 Equilibrium

We now close the model, assuming that the assembling firm of industry i sells its output to N countries which have CES demand functions

$$X_{i,n,t} = \omega_{i,n,t} X_{n,t} \left[\frac{P^*_{i,n,t}}{\mathcal{P}^*_{n,t}} \right]^{-\sigma}. \tag{12}$$

Here, $P^*_{i,n,t}$ is the price of the good in local currency (as indicated by the asterisk), $X_{n,t}$ is the country's GDP, and $\mathcal{P}^*_{n,t}$ the associated price index. Each country's demand is subject to an industry-specific demand shock $\omega_{i,n,t}$. Let $\mathcal{E}_{n,t}$ be the exchange rate in US dollars per unit of foreign currency. Then the dollar-denominated price is

$$P_{i,n,t} = \mathcal{E}_{n,t} P^*_{i,n,t}, \tag{13}$$

and since there is no pricing-to-market $P_{i,n,t} = P_{i,t}$. Letting $G_{i,t}$ denote government purchases of the good of industry i, market clearing requires that

$$X_{i,t} = G_{i,t} + \sum_n X_{i,n,t}. \tag{14}$$

An equilibrium in industry i satisfies equations (11) through (14).

We next solve for the reduced form after log-linearizing around the equilibrium in $t-1$. Before showing the equations, it is useful to define and discuss the following objects. First, the change of industry i's effective nominal exchange rate is

$$\Delta e_{i,t} := \sum_n s_{i,n,t-1} \Delta \ln \mathcal{E}_{n,t}. \tag{15}$$

In this expression $s_{i,n,t-1}$ denotes the $t-1$ sales share of industry i to country n. Notice that $\Delta e_{i,t}$ varies by industry because existing trade linkages differentially expose industries to fluctuations of a common set of currencies. The effective exchange rate change as defined in (15) takes into account that some industries sell more of their goods abroad than others. Industries that sell most of their goods domestically have a U.S. sales share near one. This share is then multiplied by $\Delta \ln \mathcal{E}_{n,t}$ which is zero for the U.S. 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 a firm or industry that sets prices in dollars such a depreciation leads to an increase in demand. Below we will use the empirical counterpart of this measure as a demand shock to test the state dependence hypothesis.

We further define

$$\Delta\xi_{i,t} := \sum_{n} s_{i,n,t-1}\Delta\ln X_{n,t} \quad \text{and} \quad \Delta\pi_{i,t} := \sum_{n} s_{i,n,t-1}\Delta\ln P^{*}_{n,t}, \tag{16}$$

which reflect changes in demand due to changes in a destination's market size and prices, respectively. Both measures rely on the same model-based weighting scheme as the exchange rate measure (15) and thus vary by industry. Finally, let $\Delta g_{i,t} := \left(G_{i,t} - G_{i,t-1}\right)/X_{i,t-1}$, that is $\Delta g_{i,t}$ is the change in government spending as a fraction of the industry's level of output at $t-1$.

Proposition 3. *The reduced form, linearized around the equilibrium in $t-1$, is*

$$\begin{aligned}\Delta\ln X_{i,t} \;&=\; \beta_e\left(u_{i,t-1}\right)\Delta e_{i,t} + \beta_\xi\left(u_{i,t-1}\right)\Delta\xi_{i,t} + \beta_\pi\left(u_{i,t-1}\right)\Delta\pi_{i,t} + \beta_g\left(u_{i,t-1}\right)\Delta g_{i,t} \\ &\quad + \beta_Q\left(u_{i,t-1}\right)\Delta\ln Q_{i,t} + \beta_{uc}\left(u_{i,t-1}\right)\Delta\ln\frac{w_{i,t}L_{i,t}}{X_{i,t}} + \omega^X_{i,t},\end{aligned} \tag{17}$$

$$\begin{aligned}\Delta\ln P_{i,t} \;&=\; \gamma_e\left(u_{i,t-1}\right)\Delta e_{i,t} + \gamma_\xi\left(u_{i,t-1}\right)\Delta\xi_{i,t} + \gamma_\pi\left(u_{i,t-1}\right)\Delta\pi_{i,t} + \gamma_g\left(u_{i,t-1}\right)\Delta g_{i,t} \\ &\quad + \gamma_Q\left(u_{i,t-1}\right)\Delta\ln Q_{i,t} + \gamma_{uc}\left(u_{i,t-1}\right)\Delta\ln\frac{w_{i,t}L_{i,t}}{X_{i,t}} + \omega^P_{i,t}.\end{aligned} \tag{18}$$

All coefficients are functions of the utilization rate (and only of the utilization rate) and $\beta_e > 0$, $\beta_\xi > 0$, $\beta_\pi > 0$, $\beta_g > 0$, $\beta_{uc} < 0$, $\beta_Q > 0$, $\gamma_e > 0$, $\gamma_\xi > 0$, $\gamma_\pi > 0$, $\gamma_g > 0$, $\gamma_{uc} > 0$, $\gamma_Q < 0$. The error terms are given by

$$\omega^X_{i,t} = \omega^X\left(u_{i,t-1}\right)\sum_n s_{i,n,t-1}\Delta\ln\omega_{i,n,t} \tag{19}$$

$$\omega^P_{i,t} = \omega^P\left(u_{i,t-1}\right)\sum_n s_{i,n,t-1}\Delta\ln\omega_{i,n,t} \tag{20}$$

where ω^X and ω^P are only functions of $u_{i,t-1}$.

4 Empirical analysis: exchange rate shocks

4.1 Data

Our preferred measures of capacity and utilization are constructed and published by the Federal Reserve Board (FRB). To obtain series for utilization, the FRB first constructs indexes of industrial production and capacity. The industrial production series are indexes of real gross output. The FRB's capacity indexes aim to capture the *sustainable maximum level of output*, that is, "the greatest level of output a plant [or industry] can maintain within the framework of a realistic work schedule after factoring in normal downtime and assuming sufficient availability of inputs to operate the capital in place".[4] The annual FRB measure of capacity is primarily based on establishment-level utilization rates from the Survey of Plant Capacity (prior to 2007), the Quarterly Survey of Plant Capacity (from 2007 onwards), and measures of capital from the Annual Survey of Manufacturers.[5] As in our model, utilization is then constructed by dividing industrial production by capacity (see equation 9).

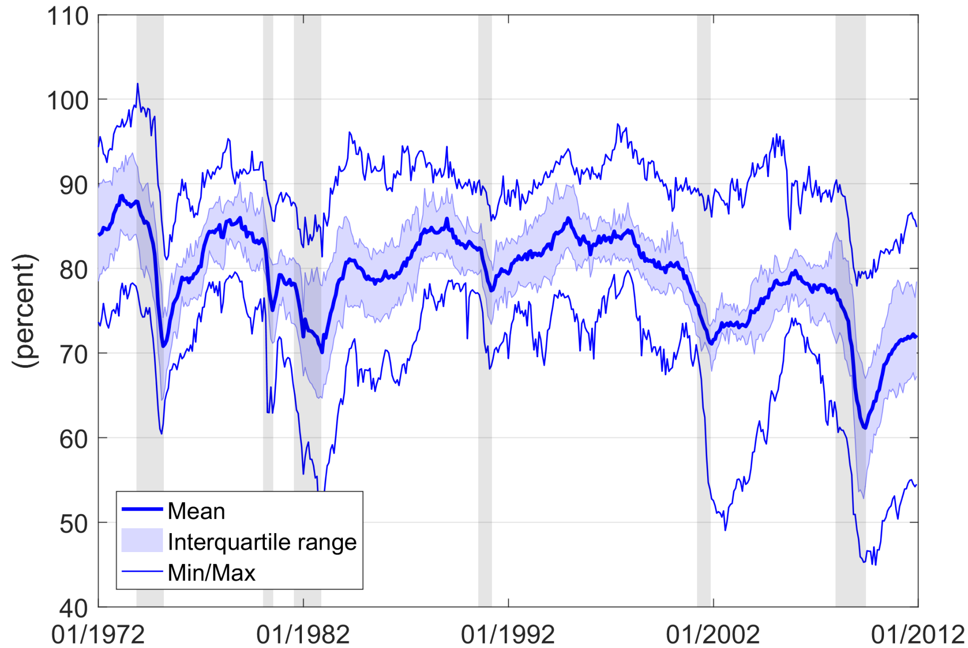
Figure 3 illustrates the capacity utilization rates of the 21 3-digit NAICS manufacturing industries in our sample. As is clear from the figure, utilization rates display significant heterogeneity both in the cross-section and over time. Capacity utilization is strongly procyclical and experiences a mild downward trend towards the end of the sample, presumably reflecting the well-documented shrinkage of the U.S. manufacturing sector. An additional salient feature which we document in Appendix table (C1) is that average utilization rates differ across industries. Since it is not clear whether these differences reflect measurement or industry-specific capacity choices we subtract from all utilization series their industry-specific means. We note, however, that this demeaning does not drive our results.

We take data on prices, sales, and input costs 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

[4]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).

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

Figure 3: Capacity utilization at the industry level



Notes: Federal reserve Board. The figure plots the mean, minimum, maximum and interquartile range of the capacity utilization series constructed using the FRB capacity utilization data and industrial production. Shaded areas represent NBER recessions.

picture of the U.S. manufacturing sector.[6] To obtain our measure of unit costs, wL/X we sum production worker wages, costs of materials, and expenditures on energy and then divide by real gross output.

We calculate the sales shares $s_{i,n,t}$ by combining the industry sales series from the NBER CES Manufacturing Industry Database with the export data available from Peter Schott's website.[7] When constructing $\Delta e_{i,t}$, $\Delta \xi_{i,t}$, and $\Delta \pi_{i,t}$ as described in (15) and (16) we limit ourselves to countries that joined the OECD prior to year 2000, but note that this choice does not drive our results. The data on exchange rates, real domestic absorption (i.e. consumption plus investment), and the associated price level are from the Penn World Tables. Our sample is annual, includes 21

[6]For a detailed description of this database, see Bartelsman and Gray (1996) and Randy Becker and Marvakov (2016).

[7]See `http://faculty.som.yale.edu/peterschott/sub_international.htm`. This data is available with SIC industry codes between 1972 and 1997, and with NAICS industry codes thereafter. We use the NBER CES SIC4 to NAICS6 concordance based on sales weights to convert the SIC codes into NAICS equivalents and then aggregate to the 3-digit NAICS level.

3-digit NAICS manufacturing industries, and ranges from 1972 to 2011.

4.2 Empirical specification and identification

To estimate equations (17) and (18), we approximate the coefficients linearly around the industry-specific mean, so that, for instance, $\beta_e(u_{i,t-1}) = \beta_e + \beta_{eu}(u_{i,t-1} - \bar{u}_i)$. (To preserve clarity and with abuse of notation, we will continue to denote the demeaned utilization measure by $u_{i,t-1}$.) Further, we add $u_{i,t-1}$ to the specification in order to obtain the conventional interpretation of the interaction terms and set government spending to zero. This yields our baseline estimating equations:

$$\begin{aligned} \Delta \ln X_{i,t} = {} & \beta_e \Delta e_{i,t} + \beta_{eu} \Delta e_{i,t} u_{i,t-1} + \beta_\xi \Delta \xi_{i,t} + \beta_{\xi u} \Delta \xi_{i,t} u_{i,t-1} \\ & + \beta_\pi \Delta \pi_{i,t} + \beta_{\pi u} \Delta \pi_{i,t} u_{i,t-1} + \beta_Q \Delta \ln Q_{i,t} + \beta_{Qu} \Delta \ln Q_{i,t} u_{i,t-1} \\ & + \beta_{uc} \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} + \beta_{ucu} \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} u_{i,t-1} + \beta_u u_{i,t-1} + \omega^X_{i,t}, \end{aligned} \tag{21}$$

$$\begin{aligned} \Delta \ln P_{i,t} = {} & \gamma_e \Delta e_{i,t} + \gamma_{eu} \Delta e_{i,t} u_{i,t-1} + \gamma_\xi \Delta \xi_{i,t} + \gamma_{\xi u} \Delta \xi_{i,t} u_{i,t-1} \\ & + \gamma_\pi \Delta \pi_{i,t} + \gamma_{\pi u} \Delta \pi_{i,t} u_{i,t-1} + \gamma_Q \Delta \ln Q_{i,t} + \gamma_{Qu} \Delta \ln Q_{i,t} u_{i,t-1} \\ & + \gamma_{uc} \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} + \gamma_{ucu} \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} u_{i,t-1} + \gamma_u u_{i,t-1} + \omega^P_{i,t}. \end{aligned} \tag{22}$$

4.2.1 The exchange rate as a demand shock

The shock we use in this section to identify the curvature of the supply curve is a (purified) change in an industry's effective exchange rate (equation 15). 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 (and 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.

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

Although there is a well-documented disconnect between exchange rate movements and other macroeconomic aggregates, it is important to acknowledge that exchange rates are endogenous variables. To the extent that the shocks driving exchange rate movements are not appropriately controlled for in our model (for instance, through $\Delta\xi_{i,t}$ and $\Delta\pi_{i,t}$), our estimates could be biased. We thus purify the exchange rate movements as follows. Let $\Delta \ln \mathcal{E}_t^{com}$ denote changes in the value of the U.S. dollar that are common to all currencies and $\Delta \ln \mathcal{E}_{n,t}^{spec}$ movements that are specific to the currency of country n. Then it is possible to write

$$\Delta \ln \mathcal{E}_{n,t} = \Delta \ln \mathcal{E}_t^{com} + \Delta \ln \mathcal{E}_{n,t}^{spec}. \tag{23}$$

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. Since these common changes could be caused by U.S. macro shocks which could also affect the manufacturing industries in our sample through other, non-exchange rate channels, we find it useful to develop a way to control for them.

To do so, denote by $n = 0, 1, ..., N$ the destination countries for manufacturing exports, letting $n = 0$ be the U.S. Since $\Delta \ln \mathcal{E}_{0,t} = 0$ for all t, we can always write $\Delta e_{i,t} = \sum_{n=0}^{N} s_{i,n,t-1} \Delta \ln \mathcal{E}_{n,t} = \sum_{n=1}^{N} s_{i,n,t-1} \Delta \ln \mathcal{E}_{n,t}$. Further, let $\bar{s}_{n,t-1}$ denote the sales share to country n for all of U.S. manufacturing, i.e. the aggregate of all industries in the sample. Then, using decomposition (23), we obtain

$$\Delta e_{i,t} = \Delta \ln \mathcal{E}_t^{com} \sum_{n=1}^{N} s_{i,n,t-1} + \sum_{n=1}^{N} \bar{s}_{n,t-1} \Delta \ln \mathcal{E}_{n,t}^{spec} + \sum_{n=1}^{N} \left(s_{i,n,t-1} - \bar{s}_{n,t-1}\right) \Delta \ln \mathcal{E}_{n,t}^{spec}. \tag{24}$$

The first term in equation (24) represents dollar movements that are common to all foreign currencies. Since industries are differentially exposed to the common component of exchange rate movements, $\Delta \ln \mathcal{E}_t^{com}$ is multiplied by the foreign sales shares $\sum_{n=1}^{N} s_{i,n,t-1} = 1 - s_{i,0,t-1}$. If global or U.S. aggregate shocks cause dollar movements relative to all foreign currencies and this variation is a concern for identification, these shocks can be perfectly controlled for by including an interaction of a set of time fixed effects with the industries' foreign sales shares in the regression. We will do so below.

The second term in equation (24) represents country-specific exchange rate movements weighted up into an effective exchange rate using the sales shares for all of U.S. manufacturing. This term does not vary by industry and can therefore be controlled for with time fixed effects.

The third and final term in equation (24) captures the remainder of the variation of the industries' effective exchange rates and constitutes the cleanest measure of our demand shock (holding unit costs constant). It weights destination-specific exchange rate movements with the deviations of sales shares from the manufacturing average. Possible confounders are now limited to shocks that 1) move destination-specific exchange rates (and not the dollar relative to all currencies), 2) differentially affect the industries in the sample in a way that is correlated with these industries' export patterns, 3) through a channel that is not the exchange rate itself, and 4) in a way that is not controlled for through $\Delta\xi_{i,t}$ and $\Delta\pi_{i,t}$ (or other controls).

It is not impossible to come up with candidates that still pose a problem for identification. Foreign monetary and financial shocks are two examples and credibly controlling for these is difficult. We therefore use an alternative demand shock to ensure that our results are not driven by these confounding factors. In the next section we will use federal defense spending to identify the curvature of the supply shock. The economic conclusions are similar, although the estimates are less precise.

4.2.2 Additional notes on identification

We briefly discuss three additional identification concerns. First, it is likely that the regression errors are correlated with utilization. Our structural approach alleviates this concern relative to other studies because it suggests a number of controls, including the change in capacity, $\Delta \ln Q_{i,t}$.[8] Yet, adding such controls does not necessarily remedy the problem fully because unobservables such as the demand shocks $\omega_{i,n,t}$ in the structural errors $\omega^X_{i,t}$ and $\omega^P_{i,t}$ could be correlated with $u_{i,t-1}$.[9] To understand this concern we ran a number of Monte-Carlo experiments and found that while the coefficients on utilization were generally biased, the coefficient on the interaction was not.

Second, and as discussed in section 3.4, variation in the interaction term $\Delta e_{i,t} u_{i,t-1}$ does not exclusively identify the curvature of the supply curve Ω'', but also the curvature in the wedge

[8]Since firms and industries with high utilization rates expand capacity, $\Delta \ln Q_{i,t}$ has an unconditional correlation with $u_{i,t-1}$ of 39%.

[9]This would be the case, for instance, if the unobserved demand shocks in changes, $\Delta \ln \omega_{i,n,t}$, were autocorrelated (see equations 19 and 20).

between marginal costs and unit costs. Proposition (2) implied that we will always understate the convexity of the supply curve, which places us on the conservative side.

Third and finally, our model is likely misspecified. Features such as price stickiness and more general specifications of heterogeneity are important features of the real world. We will consider generalizations of the model and the associated robustness exercises below.

4.3 Results

We begin with estimating of linear versions of (21) and (22). The estimates for $\Delta \ln X_{i,t}$ as the dependent variable are shown in table 1. In the absence of our model's guidance the canonical starting point is to omit capacity on the right hand side. We report the associated estimates as specification (1). All variables have the expected signs, but the coefficient on the effective exchange rate is not significant. Once we add capacity to the model (specification 2), the exchange rate response increases and becomes statistically significant at the 1 percent level. The observation that all other coefficient estimates also change (in particular the coefficient on $\Delta\pi_{i,t}$) suggests that controlling for capacity may be important. When interpreted through the lens of the model, controlling for capacity captures all relevant dynamic margins for production and price setting. While this interpretation is presumably too strong, it is likely that changes in capacity do help control for industries' low frequency dynamics or anticipation effects. Note that specification (2), which is quite parsimonious with only 5 regressors, explains about 50 percent of the variation in the data.

We next sequentially add in fixed effects to see whether doing so affects the exchange rate response. Unsurprisingly, adding industry fixed effects to account for differential growth rates changes little (specification 3). The exchange rate response shrinks somewhat when time fixed effects are also included (specification 4), but the qualitative interpretation remains the same. In specification (5) we additionally add a set of time fixed effects interacted with industries' foreign sales share to purify the exchange rate shock from potential confounders that lead to dollar movements against all foreign currencies. The fact that the coefficient estimate on $\Delta e_{i,t}$ barely changes increases our confidence that $\Delta e_{i,t}$ accurately captures changes in foreign demand after dollar depreciations rather than an omitted variable.

Table 2 presents an analogous set of estimates for $\Delta \ln P_{i,t}$ as the dependent variable. The most

Table 1: Linear model for dependent variable $\Delta \ln X_{i,t}$

Specification	(1)	(2)	(3)	(4)	(5)
$\Delta e_{i,t}$	2.14	3.51***	3.60***	2.69**	2.56*
	(1.30)	(1.09)	(1.07)	(1.08)	(1.31)
$\Delta \xi_{i,t}$	1.81***	1.57***	1.57***	-1.29	4.94
	(0.24)	(0.20)	(0.20)	(2.04)	(4.16)
$\Delta \pi_{i,t}$	0.42**	0.05	0.09	-3.78	-3.48
	(0.20)	(0.16)	(0.18)	(2.27)	(3.16)
$\Delta \ln Q_{i,t}$		0.71***	0.71***	0.79***	0.76***
		(0.07)	(0.09)	(0.08)	(0.08)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$	-0.18**	-0.10	-0.13**	-0.20**	-0.20**
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)
Observations	819	819	819	819	819
R-squared	0.355	0.498	0.509	0.641	0.677
Industry FE	no	no	yes	yes	yes
Time FE	no	no	no	yes	yes
Time FE $\times (1 - s_{i,0,t-1})$	no	no	no	no	yes

Notes: Driscoll-Kraay standard errors in parentheses. Significance levels are indicated by *** p<0.01, ** p<0.05, * p<0.1.

striking feature in this table is the importance of unit costs to explain prices. A regression of the percent change in prices on the percent change in unit costs (not reported) yields an R-squared of 0.900, just marginally below the R-squared of specification (1) in table 2. All other variables, including the exchange rate variable $\Delta e_{i,t}$, are relatively unimportant and not generally significant, independent of whether we control for capacity (specification 2) and whether we add additional fixed effects (specifications 3 to 5). The fact that even well identified shocks have little effect on prices is unfortunate, but not uncommon in the literature (see, e.g. House and Shapiro and well as Shapiro (criticism of utilization) and others).

We present the first set of state-dependent specifications with $\Delta \ln X_{i,t}$ as the dependent variable in table 3. Specification (1) shows the estimates of equation (21). Both the main effect of the effective exchange rate as well as the interaction term with lagged utilization are statistically significant. The coefficient on the interaction term is also economically very large. At the 10th percentile of the demeaned utilization rate ($u_{i,t-1} = -0.085$), a one percent depreciation of the dollar leads to a $4.69\,(= 2.79 + (-22.30) \times (-0.085))$ percent increase in the industry's quantity produced. In contrast, the same elasticity is only $1.11\,(= 2.79 + (-22.30) \times 0.075)$ at the 90th percentile ($u_{i,t-1} = 0.075$). While all our results indicate strong state-dependence, we would like

Table 2: Linear model for dependent variable $\Delta \ln P_{i,t}$

Specification	(1)	(2)	(3)	(4)	(5)
$\Delta e_{i,t}$	-0.16	-0.19	-0.19	-0.22	0.30
	(0.29)	(0.26)	(0.25)	(0.29)	(0.74)
$\Delta \xi_{i,t}$	-0.02	-0.01	-0.02	0.58	-4.02*
	(0.04)	(0.04)	(0.03)	(0.66)	(2.22)
$\Delta \pi_{i,t}$	0.12***	0.13***	0.14***	0.15	-1.19
	(0.03)	(0.03)	(0.04)	(0.63)	(1.08)
$\Delta \ln Q_{i,t}$		-0.02	0.00	0.00	0.02
		(0.03)	(0.04)	(0.04)	(0.03)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$	0.83***	0.83***	0.82***	0.87***	0.86***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Observations	819	819	819	819	819
R-squared	0.902	0.902	0.904	0.918	0.925
Industry FE	no	no	yes	yes	yes
Time FE	no	no	no	yes	yes
Time FE $\times (1 - s_{i,0,t-1})$	no	no	no	no	yes

Notes: Driscoll-Kraay standard errors in parentheses. Significance levels are indicated by *** p<0.01, ** p<0.05, * p<0.1.

caution the reader about taking these results too literally. We discuss in section X that the elasticity estimates at a particular point of the utilization distribution are somewhat sensitive to the transformation of utilization prior to estimation.

Specifications (2) to (4) of table 3 sequentially add industry fixed effects, time fixed effects, and the interaction of time fixed effects with the industries' foreign sales shares. In all cases is the coefficient on the interaction term of the exchange rate with utilization large and significant. In fact, it becomes larger as more fixed effects are added to the regression. Finally, specification (5) includes as additional controls squared terms of $u_{i,t}, \Delta e_{i,t}, \Delta \xi_{i,t}, \Delta \pi_{i,t}, \Delta \ln Q_{i,t}, \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$ as well as all possible interactions of $\Delta e_{i,t}, \Delta \xi_{i,t}, \Delta \pi_{i,t}, \Delta \ln Q_{i,t}, \Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$. Again, the coefficient of interest is essentially unchanged.

Table 4 presents an analogous set of estimates for $\Delta \ln P_{i,t}$ as the dependent variable. As in the linear specifications of table 2 the main effect of the effective exchange rate is not generally significant. However, the coefficient on the interaction term with the utilization rate is positive and significant, consistent with the predictions of a convex supply curve. Further, this result is robust to including industry fixed effects, time fixed effects, the interaction of time fixed effects with the industries' foreign sales shares, and the same set of interactions and squares as discussed above.

Table 3: State-dependent model for dependent variable $\Delta \ln X_{i,t}$

	(1)	(2)	(3)	(4)	(5)
$\Delta e_{i,t}$	2.79***	2.85***	1.27	2.02**	1.87
	(0.75)	(0.72)	(0.90)	(0.85)	(2.07)
$\Delta e_{i,t} \times u_{i,t-1}$	-22.30**	-25.44***	-24.40***	-31.35***	-30.00***
	(9.43)	(7.88)	(7.38)	(6.34)	(7.19)
$\Delta \xi_{i,t}$	1.43***	1.39***	0.52	2.05	-12.46
	(0.19)	(0.19)	(2.57)	(5.29)	(19.39)
$\Delta \xi_{i,t} \times u_{i,t-1}$	-2.27	-2.43	-2.08	-1.72	-1.29
	(1.39)	(1.45)	(1.78)	(1.20)	(1.23)
$\Delta \pi_{i,t}$	0.30	0.32	-5.49**	-6.29*	-20.64
	(0.18)	(0.20)	(2.02)	(3.39)	(14.63)
$\Delta \pi_{i,t} \times u_{i,t-1}$	-4.52**	-3.72	-3.62	-4.51	-2.98
	(2.16)	(2.34)	(3.25)	(3.15)	(2.66)
$\Delta \ln Q_{i,t}$	0.81***	0.98***	0.94***	0.92***	0.76***
	(0.06)	(0.13)	(0.13)	(0.12)	(0.15)
$\Delta \ln Q_{i,t} \times u_{i,t-1}$	-0.58	-0.56	-0.52	0.46	0.40
	(0.73)	(0.91)	(0.78)	(0.82)	(0.94)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$	-0.06	-0.10	-0.15*	-0.15*	-0.09
	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} \times u_{i,t-1}$	-0.12	-0.27	0.09	0.67	-0.36
	(1.04)	(0.90)	(0.74)	(0.73)	(0.59)
$u_{i,t-1}$	-0.04	-0.09	-0.16	-0.18**	-0.23**
	(0.09)	(0.08)	(0.10)	(0.09)	(0.08)
Observations	819	819	819	819	819
R-squared	0.554	0.569	0.679	0.714	0.735
Industry FE	no	yes	yes	yes	yes
Time FE	no	no	yes	yes	yes
Time FE $\times (1 - s_{i,0,t-1})$	no	no	no	yes	yes
Higher order controls*	no	no	no	no	yes

Notes: Driscoll-Kraay standard errors in parentheses. Significance levels are indicated by *** p<0.01, ** p<0.05, * p<0.1.

Economically, the effect is fairly small. Taking the estimates from specification (4), prices rise by $0.68\,(= 0.25 + 5.66 \times 0.075)$ in response to a one percent dollar depreciation relative to the relevant set of foreign currencies when the industry is at the 90th percentile of utilization.[10]

[10] Notice that since our effective exchange rate measure $\Delta e_{i,t}$ places a large weight on the domestic sales share, its standard deviation is small (approximately 0.5 percent). A one percent shock to the effective exchange rate is therefore a very large shock.

Table 4: State-dependent model for dependent variable $\Delta \ln P_{i,t}$

	(1)	(2)	(3)	(4)	(5)
$\Delta e_{i,t}$	-0.12	-0.13	-0.13	0.25	-1.69*
	(0.24)	(0.23)	(0.21)	(0.74)	(0.93)
$\Delta e_{i,t} \times u_{i,t-1}$	4.89**	4.53**	5.99***	5.66***	5.10**
	(1.74)	(1.78)	(1.97)	(1.66)	(1.93)
$\Delta \xi_{i,t}$	0.03	0.03	0.54	-3.48*	-6.35
	(0.05)	(0.05)	(0.58)	(1.85)	(4.68)
$\Delta \xi_{i,t} \times u_{i,t-1}$	1.01***	1.03***	0.73*	0.62*	0.29
	(0.30)	(0.31)	(0.37)	(0.35)	(0.45)
$\Delta \pi_{i,t}$	0.09**	0.09**	0.26	-1.45	-2.04
	(0.04)	(0.04)	(0.61)	(1.14)	(2.33)
$\Delta \pi_{i,t} \times u_{i,t-1}$	0.74	0.87	0.58	0.57	0.28
	(0.79)	(0.70)	(0.86)	(0.81)	(0.89)
$\Delta \ln Q_{i,t}$	-0.01	0.00	0.00	0.02	0.01
	(0.02)	(0.03)	(0.03)	(0.03)	(0.06)
$\Delta \ln Q_{i,t} \times u_{i,t-1}$	-0.47	-0.39	-0.47	-0.45	-0.65
	(0.33)	(0.31)	(0.31)	(0.29)	(0.52)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$	0.83***	0.83***	0.87***	0.87***	0.82***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
$\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}} \times u_{i,t-1}$	0.46	0.44	0.55	0.45	0.45
	(0.33)	(0.34)	(0.37)	(0.31)	(0.38)
$u_{i,t-1}$	-0.06*	-0.07**	-0.04	-0.04	-0.01
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Observations	819	819	819	819	819
R-squared	0.906	0.908	0.921	0.927	0.932
Industry FE	no	yes	yes	yes	yes
Time FE	no	no	yes	yes	yes
Time FE $\times (1 - s_{i,0,t-1})$	no	no	no	yes	yes
Higher order controls*	no	no	no	no	yes

Notes: Driscoll-Kraay standard errors in parentheses. Significance levels are indicated by *** p<0.01, ** p<0.05, * p<0.1.

5 Empirical analysis: government spending shocks

In this section we expand the empirical analysis to government spending shocks to ensure that our results are robust. Many of the details are deferred to the appendix.

5.1 Data

Defense spending series at the industry level are constructed from the military prime contract files and USAspending.gov. The military prime contract files include information on all military prime contracts with values above the minimum threshold of \$10,000 up to 1983 and \$25,000 thereafter.

They can be downloaded for the period from 1966 to 2003 from the U.S. National Archives. We complement this data source with data from USAspending.gov, a government website dedicated to promoting transparency of federal spending. The data from USAspending.gov is available from 2000 onwards. A comparison of the two data sources for the overlapping years from 2000 to 2003 reveals only negligible differences. To assign defense spending to NAICS 3-digit industries, we construct several concordances which we discuss in Appendix X. The remaining data for the analysis are as described in the previous section.

5.2 Empirical specification and identification

We estimate Jordà (2005) local projections of the form

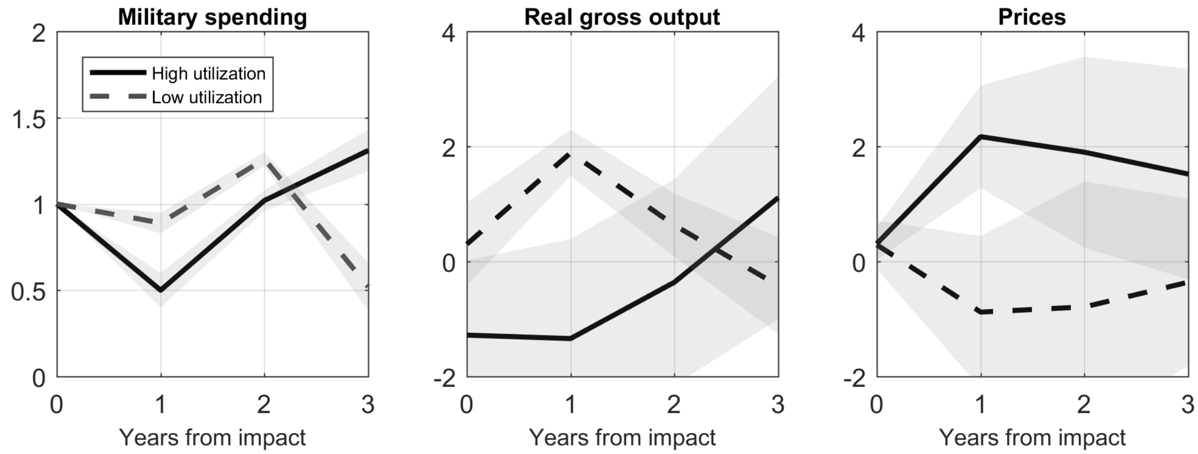
$$\begin{aligned} \frac{X_{i,t+h} - X_{i,t-1}}{X_{i,t-1}} &= \alpha_h^L \left(1 - F_{i,t-1}\right) \Delta g_{i,t} + \alpha_h^H F_{i,t-1} \Delta g_{i,t} \\ &\quad + \delta_h Z_{i,t} + \mu_{i,h} + \zeta_{t,h} + \left(1 - s_{i,0,t-1}\right) \chi_{t,h} + \varepsilon_{i,t+h}, \end{aligned} \tag{25}$$

where $F_{i,t} = \frac{\exp(\gamma \cdot u_{i,t})}{1+\exp(\gamma \cdot u_{i,t})}$ and $u_{i,t}$ is the *demeaned* utilization rate. We set γ to 25 which makes this inverse-logit transformation essentially identical to the empirical cumulative distribution function, and hence $F_{i,t}$ can be interpreted as a percentile of the distribution of demeaned utilization rates. The vector of controls $Z_{i,t}$ includes interactions of $\Delta g_{i,t-1}$, $\Delta \xi_{i,t}$, ,$\Delta \pi_{i,t}$, $\Delta \ln Q_{i,t}$, and $\Delta \ln \frac{w_{i,t} L_{i,t}}{X_{i,t}}$ with $1 - F_{i,t-1}$ and $F_{i,t-1}$ as well as the main effect of $F_{i,t-1}$. Because $\Delta g_{i,t} = \frac{G_{i,t} - G_{i,t-1}}{X_{i,t-1}}$, the impulse response coefficients α_h^L for low initial utilization rates and α_h^H for high initial utilization rates are interpreted as dollar-for-dollar changes. We continue to include industry fixed effects, time fixed effects, and the interaction of time fixed effects with the foreign sales share. Adding the effective exchange rate to the specification makes little difference.

We construct Bartik-type instruments for defense spending at the industry level similar to Nakamura and Steinsson (2014) and Boehm (2016). Let G_t and X_t denote real defense spending and real gross output of all of U.S. manufacturing. The instrument is then the average pre-sample share of government purchases multiplied by change in defense spending on all of manufacturing. Formally, the instrument for industry i is

$$\frac{1}{6} \sum_{\tau=1966}^{1971} \frac{G_{i,\tau}}{X_{i,\tau}} \times \frac{G_t - G_{t-1}}{X_{t-1}}, \quad t = 1972, ..., 2011.$$

Figure 4: Impulse response functions for a government spending shock



Notes: The impulse response coefficients are obtained via local projections using specification (25) with the controls $Z_{i,t}$ as described in the text. The shaded areas represent one standard error bands, based on standard errors that are clustered at the industry level. Impulse response functions for high utilization correspond to setting $F_{i,t-1} = 1$ and low utilization to setting $F_{i,t-1} = 0$.

As in Nakamura and Steinsson (2014) and Boehm (2016) all diagnostics suggest that the instruments are strong. We report them in Appendix XXX.

5.3 Results

Figure 4 shows the impulse response functions for defense spending, real gross output, and prices. A one dollar increase in defense spending is followed by additional spending in subsequent years. Over the course of the sample period, increases in spending that started at low initial levels of utilization were followed by somewhat higher spending one and two years after the shock. Cumulative spending over the first four years, including the impact year, is comparable: 3.83$ for high initial utilization rates versus 3.67$ for low initial utilization rates.

Industry gross output rises only when the initial capacity utilization rate is low. This is consistent with the earlier findings that were based on exchange rate variation and suggests that our finding of state-dependent responses to demand shocks is robust. Notice, however, that the standard errors are very wide, in particular for high initial utilization rates. (The shaded areas represent one standard deviation error bars.) The estimated multiplier for industry gross output when the initial utilization rate is low is 1.16 at a time horizon of 2 years and 0.9 at a time horizon of 3 years. The multipliers for high initial utilization rates are negative, but the standard errors are too large to reject the null hypothesis of equal multipliers at high and low initial levels of capacity utilization.

As the rightmost panel of figure 4 shows, prices rise when the spending shock occurs at high initial utilization rates and the response is significant at the 5 percent level one year after the shock. This impulse response is estimated using specification (25) after replacing the dependent variable with the percent change in prices. The response is insignificant if the spending shock occurs at low levels of capacity utilization. Note that these findings are again consistent with our earlier findings.

We discuss in the Appendix several robustness checks, including a sample split depending on whether the initial utilization rate is above or below its median. The results are very similar and reported in figure C1.

6 Conclusion

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

To be completed.

References

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

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

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

———. 2013b. "Output Spillovers from Fiscal Policy." *American Economic Review* 103 (3):141–146. URL `https://ideas.repec.org/a/aea/aecrev/v103y2013i3p141-46.html`.

Bachmann, Ruediger and Giuseppe Moscarini. 2011. "Business Cycles and Endogenous Uncertainty." Tech. rep.

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

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

Berger, David and Joseph Vavra. 2017. "Shocks vs. Responsiveness: What Drives Time-Varying Dispersion?" 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.

Boehm, Christoph E. 2016. "Government Spending and Durable Goods." Tech. rep.

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

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

Fagnart, Jean-Francois, Omar Licandro, and Henri R. Sneessens. 1997. "Capacity utilization and market power." *Journal of Economic Dynamics and Control* 22 (1):123–140. URL `https://ideas.repec.org/a/eee/dyncon/v22y1997i1p123-140.html`.

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

Fazzari, Steven M., James Morley, and Irina Panovska. 2015. "State-dependent effects of fiscal policy." *Studies in Nonlinear Dynamics & Econometrics* 19 (3):285–315. URL `https://ideas.repec.org/a/bpj/sndecm/v19y2015i3p285-315n5.html`.

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.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1):161–182.

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

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

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

Nakamura, Emi and Jn Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104 (3):753–92.

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

Peersman, Gert and Frank Smets. 2005. "The Industry Effects of Monetary Policy in the Euro Area." *Economic Journal* 115 (503):319–342. URL `https://ideas.repec.org/a/ecj/econjl/v115y2005i503p319-342.html`.

Ramey, Valerie A. and Sarah Zubairy. 2014. "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data." Nber working papers, National Bureau of Economic Research, Inc.

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

Santoro, Emiliano, Ivan Petrella, Damjan Pfajfar, and Edoardo Gaffeo. 2014. "Loss aversion and the asymmetric transmission of monetary policy." *Journal of Monetary Economics* 68 (C):19–36.

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

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

Tenreyro, Silvana and Gregory Thwaites. 2016. "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions." *American Economic Journal: Macroeconomics* 8 (4):43–74.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *The Quarterly Journal of Economics* 129 (1):215–258.

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

A Appendix: Model Extensions

A.1 Fagnart et al.

A.2 Intermediate goods producers

A.2.1 Production Function

To introduce the notion of capacity constraints our framework departs from standard theory by assuming that a bundle of factors k_t is fixed in the short run. k_t could be structures, equipment, or even specialized workers. Further, the firm has to decide ex-ante the maximum of variable factors, b_t, that it can employ (or process) in the short run. Since variable factors include primarily production workers and intermediate inputs, b_t has the natural interpretation of workstations or processing capacity of intermediates.[11]

We assume that the *production capacity* of a firm is given by

$$q_t = z_t F\left(k_t, b_t\right) \tag{A1}$$

where z_t is productivity and F is constant returns to scale. The firm's actual production x_t is

$$x_t = q_t \frac{l_t}{b_t} = z_t F\left(\kappa_t, 1\right) l_t \tag{A2}$$

where $\kappa_t = \frac{k_t}{b_t}$ and l_t represents the firm's choice of a bundle of factors that are variable in the short run. Production x_t is limited in the short run because the variable factors l_t cannot exceed b_t, that is $l_t \leq b_t$. Note that the marginal product of the variable factors, $z_t F\left(\kappa_t, 1\right)$, is constant in the short run and determined by z_t and κ_t. Letting w_t denote the price of input bundle l_t, short run marginal costs are

$$mc_t = \frac{w_t}{z_t F\left(\kappa_t, 1\right)}. \tag{A3}$$

A.3 Aggregating firm

The perfectly competitive aggregating firm uses a constant returns to scale (CES) production function with elasticity of substitution θ. Unlike the standard model, however, the aggregating firm's suppliers are subject to capacity constraints. Whenever a supplier's production is limited by its capacity the aggregating firm is constrained in the factor market.

The aggregating firm maximizes

$$\max P_t X_t - \int_0^1 p_t\left(j\right) x_t\left(j\right) dj \tag{A4}$$

subject to the capacity constraints of the intermediate suppliers

$$x_t\left(j\right) \leq q_t\left(j\right) \ \forall j \tag{A5}$$

[11] This paper emphasizes a technological interpretation of capacity constraint. An alternative is that firms do not find it optimal to produce above the level of capacity.

and its production technology

$$X_t = \left(\int_0^1 v_t(j)^{\frac{1}{\theta}} x_t(j)^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}}. \tag{A6}$$

Here, P_t is the price index, X_t is output, and $p_t(j)$ and $x_t(j)$ are the price and quantity of variety j. The production function of the aggregate bundle features variety specific shocks $\nu_t(j)$ which represent the importance of the associated intermediate input. For simplicity, we assume that $v_t(j)$ are i.i.d. shocks with cumulative distribution function $G(v)$ and unit mean.

To maintain tractable aggregation, we also make a set of timing assumptions that will guarantee that $p_t(j) = p_t$ and $q_t(j) = q_t$ $\forall j$. We will discuss these assumptions below. Imposing this symmetry implies that the factor demand functions are

$$x_t = \begin{cases} v_t X_t \left[\frac{p_t}{P_t}\right]^{-\theta} & \text{if} \quad v_t < \bar{v}_t \\ q_t & \phantom{\text{if}} \quad v_t \geq \bar{v}_t \end{cases} \tag{A7}$$

for all j, where the threshold variety $\bar{v}_t$ above which the supply of intermediates is rationed satisfies

$$\bar{v}_t = \frac{q_t}{X_t \left[\frac{p_t}{P_t}\right]^{-\theta}}. \tag{A8}$$

The price index of the aggregating firm is (in logs)

$$\ln(P_t) = \ln(p_t) + \Omega_t. \tag{A9}$$

where

$$\Omega_t = \frac{1}{1-\theta} \ln \left(\int_0^1 v(j) \left[1 + \frac{\lambda_t(j)}{p_t}\right]^{1-\theta} dj \right).$$

In this equation $\lambda_t(j)$ are the multipliers on the capacity constraints. Note that if all suppliers are unconstrained, $\lambda_t(j) = 0$ for all j, and the price index P_t equals the (common) price of intermediates p_t. However, when the supplier's capacity falls short of the quantity demanded, the aggregating firm cannot equate the marginal product of the intermediate to its price. Instead, the marginal product exceeds its price and the firm begins to earn profits.

In order to raise production the aggregating firm must now purchase input varieties from suppliers that are not constrained and these varieties have lower marginal products. As a result marginal costs rise. We call this increment of the price index above the price of intermediates the *rationing wedge* Ω_t. Whenever some input varieties are rationed, this wedge will be positive. It is easy to show that Ω_t is only a function of the threshold variety $\bar{v}_t$,

$$\Omega_t = \frac{1}{1-\theta} \ln \left(\int_0^{\bar{v}_t} v dG(v) + \bar{v}_t^{\frac{\theta-1}{\theta}} \int_{\bar{v}_t}^{\infty} v^{\frac{1}{\theta}} dG(v) \right). \tag{A10}$$

Similarly, production X_t of the industry can be written as a function of capacity q_t and $\bar{v}_t$,

$$X(q_t, \bar{v}_t) := q_t \left(\left(\frac{1}{\bar{v}_t} \right)^{\frac{\theta-1}{\theta}} \int_0^{\bar{v}_t} v dG(v) + \int_{\bar{v}_t}^{\infty} v^{\frac{1}{\theta}} dG(v) \right)^{\frac{\theta}{\theta-1}}. \tag{A11}$$

Capacity utilization We define the industry's utilization rate as the ratio of actual production to the hypothetical level of output that would be attainable if every intermediate firm produced at full capacity, that is

$$u_t := \frac{X(q_t, \bar{v}_t)}{\lim_{\bar{v}_t \to 0} X(q_t, \bar{v}_t)}. \tag{A12}$$

This definition has several attractive properties. First, it is constructed very similar to its empirical counterpart. For example, it is possible to express u_t as an (appropriately constructed) average of firms' idiosyncratic utilization rates $u_t(j) := \frac{x(j)}{q_t} = \frac{p_t x_t(j)}{p_t q_t}$. The surveys of plant capacity that we will use below construct utilization measures by asking respondents to state the market value of their actual production $p_t x_t(j)$ and the market value of their full capacity production $p_t q_t$. Further, since $\lim_{\bar{v}_t \to 0} X(q_t, \bar{v}_t) \propto q_t$ it follows that $u_t \propto \frac{X_t}{q_t}$, that is, utilization is proportional to actual production divided by capacity. This corresponds closely to the utilization series that the Federal Reserve constructs by dividing an industrial production index by an index of production capacity.

Lemma 2. *The utilization rate as defined in (9) has the following properties:*

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

The lemma highlights that the aggregate utilization rate is only a function of the threshold value $\bar{v}_t$ above which intermediate input suppliers produce at maximum capacity. The utilization rate approaches zero if no intermediate input supplier is producing at full capacity. It tends to one if all intermediate input suppliers become capacity constrained. Further, u is decreasing everywhere, and this implies that u is invertible and we can write $\bar{v}(u_t)$. We will make extensive use of this feature, both for the remainder of the theoretical analysis and when taking the model to the data. One immediate application is that we can write the rationing wedge as a function of utilization: $\Omega_t = \Omega(u_t)$. Figure A.3 illustrates a numerical example of the variation of u with $\bar{v}$.

The rationing wedge With Lemma 1 in hand we can now summarize the properties of $\Omega(u_t)$.

Proposition 4. *Under minor assumptions on* G

1. $\Omega'(u) \geq 0$
2. $\lim_{u \to 0} \Omega(u) = 0$, $\lim_{u \to 1} \Omega(u) = \infty$
3. $\lim_{u \to 0} \Omega'(u) = 0$, $\lim_{u \to 1} \Omega'(u) = \infty$

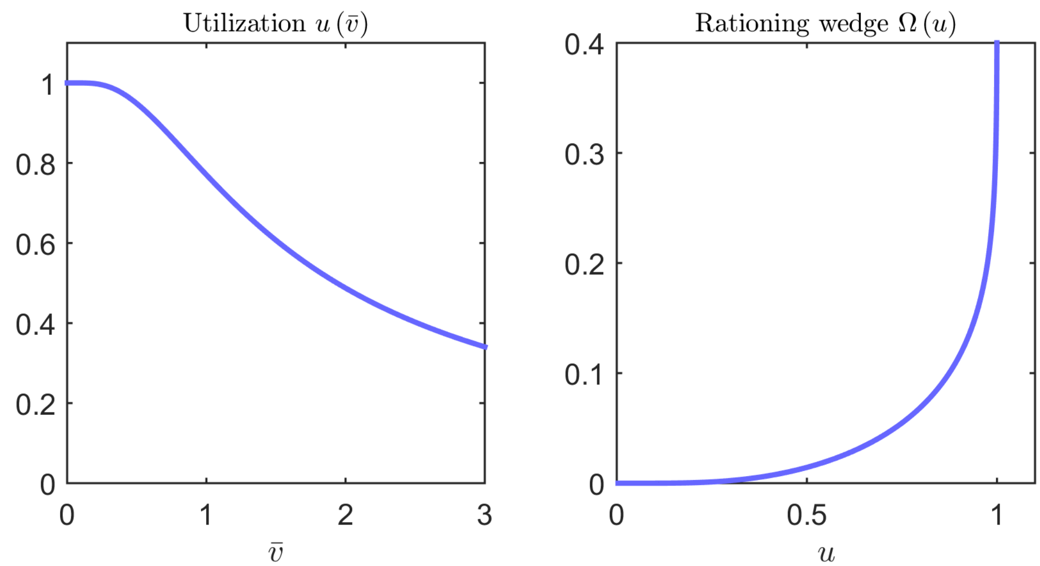


Figure A1: Utilization and the Rationing Wedge

4. *Without further restrictions on G, the sign of $\Omega''(u)$ is generally ambiguous.*

The rationing wedge is increasing in utilization everywhere. As an increasing number of suppliers become constrained, the input allocation becomes worse and the wedge widens. When the utilization rate approaches one, all suppliers become constrained and the wedge and its derivative tend to infinity. Conversely, when the utilization rate tends to zero, intermediate input suppliers are no longer capacity constrained. As a result both the rationing wedge and its derivative tend to zero. While $\Omega(u)$ is convex everywhere for many choices of G, it is possible to construct examples in which the wedge is locally concave. We conclude that the rationing wedge is a potential source of convexity of the supply curve, but whether this is so in the relevant range of utilization remains an empirical question. The right panel of Figure A.3 illustrates a numerical example of the rationing wedge.

A.4 Intermediate Goods Firms

Before turning to the intermediate goods producers, we describe the timing assumptions we make. As noted before the objective of these assumptions is to preserve simple aggregation properties of the model. Over the course of a period, firms

1. choose their fixed factor bundle k_t and the maximum processing capacity b_t
2. learn the aggregate productivity draw z_t
3. choose their price p_t
4. learn their idiosyncratic demand shock v_t
5. choose the variable factor bundle l_t and produce.

Since the choices of k_t and b_t are secondary for our empirical analysis, we relegate them to Appendix A. For the remainder of the paper it is sufficient to know that k_t and b_t have been chosen optimally and are fixed until the end of the period.

Having learned the aggregate productivity draw z_t, firms choose their price p_t. They maximize expected profits

$$E_v\left[(p_t - mc_t)\, x_t\right] \tag{A13}$$

subject to the demand curve (2). To understand the firms' incentives, it is useful to first consider the expected quantity sold,

$$E_v\left[x_t\right] = X_t \left[\frac{p_t}{P_t}\right]^{-\theta} \int_0^{\bar{v}_t} v dG\left(v\right) + q_t \int_{\bar{v}_t}^{\infty} dG\left(v\right).$$

When setting their price, firms take into account that they become capacity constrained when demand materializes sufficiently high. Conditional on becoming constrained, there are no costs of choosing higher prices because doing so does not reduce the quantity sold. As a result, the relevant notion of the demand elasticity for the firm is

$$-\frac{\partial \ln E_v\left[x_t\right]}{\partial \ln p_t} = \theta \frac{\int_0^{\bar{v}(u_t)} v dG\left(v\right)}{\int_0^{\bar{v}(u_t)} v dG\left(v\right) + \bar{v}\left(u_t\right) \int_{\bar{v}(u_t)}^{\infty} dG\left(v\right)} =: \tilde{\theta}\left(u_t\right) \tag{A14}$$

We call this elasticity the effective demand elasticity. Clearly, $\tilde{\theta}\left(u_t\right) \in [0, \theta]$. Further, since the industry-wide utilization measure u is a sufficient statistic for the firm to predict whether it will be capacity constrained, $\tilde{\theta}$ is a function of u_t (and only of u_t). It is also easy to verify that under minor regularity conditions on G, $\lim_{u \to 0} \tilde{\theta}\left(u_t\right) = \theta$: As the industry's utilization rate approaches zero and the probability of being capacity constrained tends to zero and $\tilde{\theta}$ approaches the true demand elasticity. Similarly, $\lim_{u \to 1} \tilde{\theta}\left(u_t\right) = 0$.

It is not generally true that $\tilde{\theta}$ is decreasing everywhere since two competing effects govern the sign of the derivative of $\tilde{\theta}$ and either one can dominate. First, as utilization increases, the probability of becoming capacity constrained rises. This *rationing effect* reduces $\tilde{\theta}$ because for constrained suppliers the quantity does not fall when they raise the price by one marginal unit.

Second, there is a *composition effect*. To see this, suppose that capacity q_t is fixed and recall that $u_t \propto \frac{X_t}{q_t}$. Since a higher utilization rate requires higher output of the industry, the assembling firm must increase demand from suppliers that are not rationed. In expectation, this raises the quantity of output for which the demand curve is downward sloping. As a result $\tilde{\theta}$ rises.

For many choices of G, the rationing effect dominates the composition effect and $\tilde{\theta}'$ is negative everywhere. In Appendix B we present an example where this is not the case. The left panel of A2 shows a numerical example of $\tilde{\theta}$ and Lemma 3 summarizes its properties

Lemma 3. *Under minor regularity conditions on G, $\tilde{\theta}\left(u_t\right) \in [0, \theta]$ satisfies the following properties*

1. $\lim_{u \to 0} \tilde{\theta}\left(u_t\right) = \theta$, $\lim_{u \to 1} \tilde{\theta}\left(u_t\right) = 0$
2. $\lim_{u \to 0} \tilde{\theta}'\left(u_t\right) = 0$
3. *The sign of $\tilde{\theta}'$ is generally ambiguous.*

The optimal price choice implies that

$$p_t = \mathcal{M}(u_t)\, mc_t, \tag{A15}$$

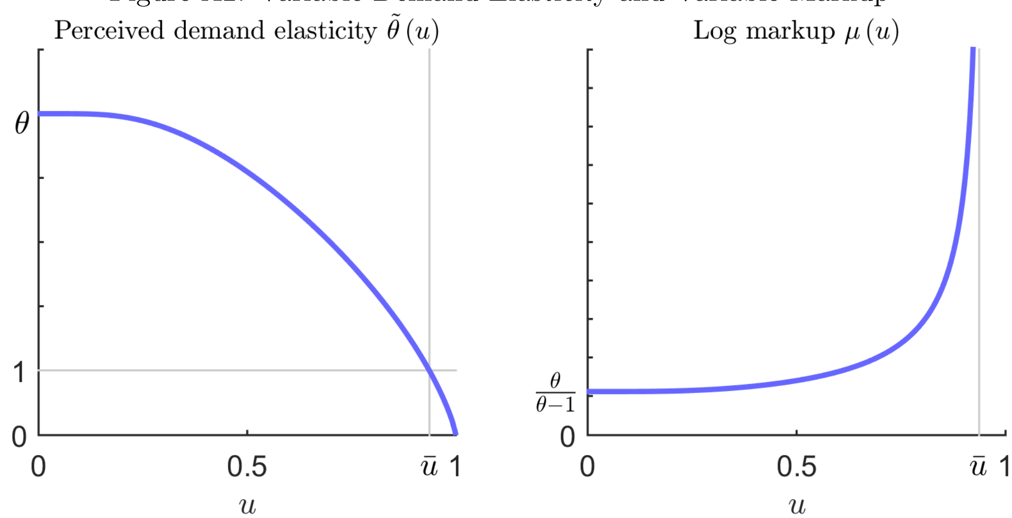
where $\mathcal{M}(u_t) = \frac{\tilde{\theta}(u_t)}{\tilde{\theta}(u_t)-1}$ and $\tilde{\theta}(u_t)$ is given by (A14). Clearly, the effective demand elasticity cannot fall below one because the markup would not be defined. The properties of $\mathcal{M}(u_t)$ follow immediately from Lemma 3.

Proposition 5. *Let* $\bar{u} = \sup\left\{u : \tilde{\theta}(u) = 1\right\}$. *Then*

1. $\lim_{u\to 0} \mathcal{M}(u) = \frac{\theta}{\theta-1}$, $\lim_{u\uparrow \bar{u}} \mathcal{M}(u) = \infty$
2. $\lim_{u\to 0} \mathcal{M}'(u) = 0$ *and* $\lim_{u\uparrow \bar{u}} \mathcal{M}'(u) = \infty$
3. *The markup may not be increasing in* u *everywhere*

As the utilization rate approaches zero, the markup tends to $\frac{\theta}{\theta-1}$–the value that would prevail in the absence of capacity constraints. Further, as the utilization rate tends to $\bar{u}$, the markup as well as its slope go to infinity. Note further, that if the effective demand elasticity $\tilde{\theta}$ is not monotonic, the markup inherits this property. The intuition is the following. If utilization rises, more firms are constrained. This raises the markup. However, this effect can locally be dominated by the fact that the aggregating firm purchases more from firms that are not rationed. Since in these states of the world the optimal markups are lower this composition effect reduces $\mathcal{M}(u)$. Appendix *B* discusses this possibility further. The right panel of Figure A2 illustrates a numerical example of the log markup $\mu(u) = \ln \mathcal{M}(u)$.

Figure A2: Variable Demand Elasticity and Variable Markup



A.5 Equilibrium

We can now put the individual pieces together and write the supply curve of the industry as a function of the markups, the rationing wedge and marginal costs

$$\ln P_t = \mu(u_t) + \Omega(u_t) + \ln mc_t. \tag{A16}$$

Both the (log) markup and the rationing wedge are mechanisms that potentially lead to a convex supply curve.

We now close the model, assuming that the assembling firm of industry i sells its output to N countries which have CES demand functions

$$X_{i,n,t} = \omega_{i,n,t} X_{n,t} \left[\frac{P^*_{i,n,t}}{\mathcal{P}^*_{n,t}}\right]^{-\sigma}. \tag{A17}$$

Here, $P^*_{i,n,t}$ is the price of the good in local currency, $X_{n,t}$ is the country's GDP, and $\mathcal{P}^*_{n,t}$ the associated price index. Each country has a possible industry-specific demand shock $\omega_{i,n,t}$. Let $\mathcal{E}_{n,t}$ be the exchange rate in US dollars per unit of foreign currency. Then the dollar-denominated price is

$$P_{i,n,t} = \mathcal{E}_{n,t} P^*_{i,n,t}. \tag{A18}$$

Finally, market clearing requires that

$$X_{i,t} = \sum_n X_{i,n,t}. \tag{A19}$$

A.6 Empirical specifications

Proposition 6. *The reduced form of the industry's quantity, linearized around its equilibrium in $t-1$ is*

$$\Delta \ln X_{i,t} = \beta_e(u_{t-1}) e_{i,t} + \beta_\pi(u_{t-1}) \pi_{i,t} + \beta_\xi(u_{t-1}) \xi_{i,t} + \beta_q(u_{t-1}) \Delta \ln q_{i,t} + \beta_{mc}(u_{t-1}) \Delta \ln mc_{i,t} + \omega^X_{i,t}$$

where

$$\beta_e > 0, \beta_\pi > 0, \beta_\xi > 0, \beta_q > 0, \beta_{mc} < 0.$$

Further

$$e_{i,t} = \sum_n s_{i,n,t-1} \Delta \ln \mathcal{E}_{n,t}, \quad \pi_{i,t} = \sum_n s_{i,n,t-1} \Delta \ln \mathcal{P}^*_{n,t}, \quad \xi_{i,t} = \sum_n s_{i,n,t-1} \Delta \ln X_{n,t}$$

and $s_{i,n,t-1}$ are sales shares to the respective counties in $t-1$, and

$$\omega^X_{i,t} = \Gamma(u_{t-1}) \sum_n s_{i,n,t-1} \Delta \ln \omega_{i,n,t}$$

$\Gamma(u_{t-1})$ is only a function of u_{t-1}. Further, if the supply curve is convex, then

$$\beta'_e < 0,\ \beta'_\pi < 0,\ \beta'_\xi < 0,\ \beta'_q > 0,\ \beta'_{mc} > 0$$

And now the price

Proposition 7. *The reduced form of the price, linearized around its equilibrium in $t-1$ is given by*

$$\Delta \ln P_{i,t} = \gamma_e\left(u_{t-1}\right) e_{i,t} + \gamma_\pi\left(u_{t-1}\right) \pi_{i,t} + \gamma_\xi\left(u_{t-1}\right) \xi_{i,t} + \gamma_q\left(u_{t-1}\right) \Delta \ln q_{i,t} + \gamma_{mc}\left(u_{t-1}\right) \Delta \ln mc_{i,t} + \omega^P_{i,t}$$

where

$$\gamma_e > 0,\ \gamma_\pi > 0,\ \gamma_\xi > 0,\ \gamma_q < 0,\ \gamma_{mc} > 0,$$

$e_{i,t}$*,* $\pi_{i,t}$*,* $\xi_{i,t}$*, and* $s_{i,n,t-1}$ *are defined as in the earlier proposition. Further,*

$$\omega^P_{i,t} = \Xi\left(u_{t-1}\right) \sum_n s_{i,n,t-1} \Delta \ln \omega_{i,n,t}$$

and $\Xi\left(u_{t-1}\right)$ *is only a function of* u_{t-1}*. If the supply curve is convex, then*

$$\gamma'_e > 0,\ \gamma'_\pi > 0,\ \gamma_\xi > 0,\ \gamma'_q < 0,\ \gamma'_{mc} < 0$$

Discussion... (to be completed)
Threats to identification

1. $\omega_{i,t}$ may be correlated with $q_{i,t}$
2. Average costs may not equal marginal costs
3. Measurement error
4. Everything that is not in the model

A.7 Taking the model to the data

There are two ways to take the model to the data. First, we estimate the reduced form. Second, we test whether the effective demand elasticity is a function of u, as the model predicts:

$$\Delta \ln X_{i,n,t} = \beta u_{i,t-1} \Delta \ln \mathcal{E}_{n,t} + \gamma_{i,t} + \delta_{n,t} + \varepsilon_{i,n,t} \tag{A20}$$

B Data Appendix

B.1 Sample and data sources

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

The sales shares $s_{i,n,t}$ 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. Sales to other countries, not included in the list, are counted as sales to the U.S.

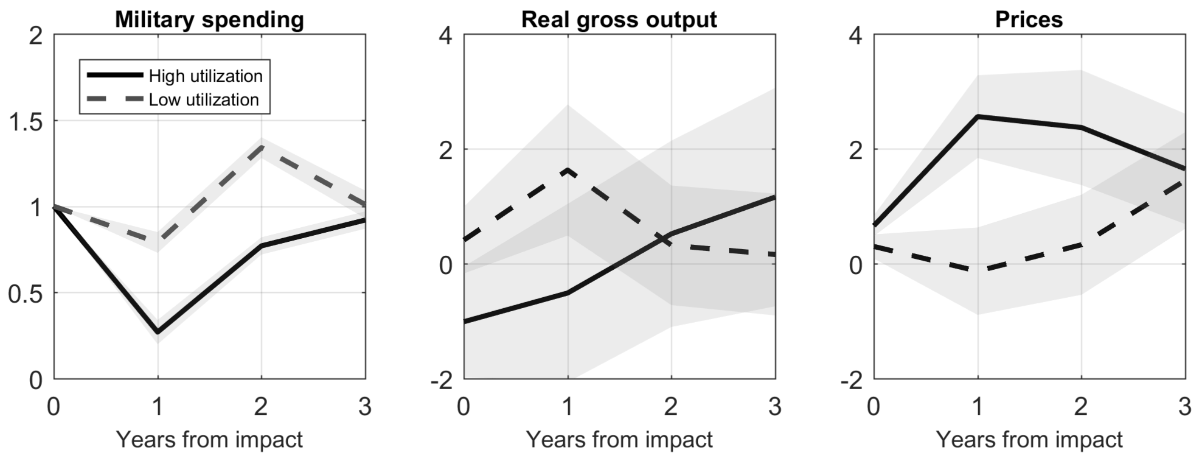
C Appendix: Additional Results

Table C1: Summary Statistics of Utilization Rates by 3-digit NAICS Manufacturing Industries

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

Source: Federal Reserve Board

Figure C1: Impulse response functions for a government spending shock



Notes: The impulse response coefficients are obtained via local projections using specification (XXX). The shaded areas represent one standard error bands, based on standard errors that are clustered at the industry level.