Journal of Economics & Management Strategy

A Map of Markups: Why We Observe Mixed Behaviors of Markups

Seong-Hoon Kim
Department of Economics
Yonsei University - Wonju
1 Yonseidae-gil, Wonju-si, Kangwon-do, 26493, Korea
s.kim.econ@gmail.com

Seongman Moon
Department of Economics and Research Institute of Big Data Business
Chonbuk National University
567 Baekje-daero, Deokjin-gu, Jeonju-si Jeollabuk-do 54896 Korea
nopasanada0501@gmail.com

This paper proposes an explanation for mixed evidence on the behaviors of markups. The key mechanism consists of two complementary channels through which firms handle uninsurable business losses. One channel is based on cost-compensating motive, by which firms raise prices to reflect higher losses stochastically associated with higher output levels. The other channel is based on loss-balancing motive, by which firms lower prices to countervail higher losses stochastically associated with higher output levels. The relative responsiveness of the two channels to a shock depends on each firm's fundamental characteristics and leads to a sharp division of markup cyclicality across sectors.

> How markups move, in response to what, and why, is almost *terra incognita* for macro. (. . .) [W]e are a long way from having either a clear picture or convincing theories, and this is clearly an area where research is urgently needed. — Blanchard (2009, p. 220)

1. Introduction

Economists have studied the behavior of markups first because it illuminates the characteristics of market structure and second because it clues to the nature of business cycles. For example, theories based on procyclical competition at the industry level predict markups to fall during booms and thus leave room for monetary expansion to have less pressure on inflation. Theories that predict markups to rise during booms would have different implications on the nature of business cycles and policy consequences. Data, however, show mixed evidence on the behavior of markups. For example, markups in Textiles (SIC code 22) and in Apparel (SIC code 23) behave differently over the U.S. business cycle (Rotemberg and Woodford, 1991, table 8.b; p. 109). Basu and Fernald (1997), Gopinath et al. (2011), and De Loecker and Warzynski (2012) have also documented sectoral and locational differences in markups. The aim of this paper is to add an

We thank Yongsung Chang and seminar participants at the 2016 Korean Econometric Society meeting for helpful comments and data. We are responsible for all errors. Corresponding Author: Seongman Moon

Journal of Economics & Management Strategy, Volume 26, Number 2, Summer 2017, 529–553

explanation for how and why the markups may behave differently across sectors thereby leading to unclear and mixed cyclicality as observed in data.

Our theory of markups takes the presence of uninsurable business losses to the fore of story. Asset markets are incomplete over goods market outcomes: General Motors safeguards its business from losses due to various natural disasters but cannot prevent losses incurred through normal business operation. The Real Greek, a small restaurant in London, holds similar insurance policies. The restaurant loss-proofs its business against unprecedented flood but is exposed to unprecedented customer visits. The managers will then take their uninsurable business losses into account when making pricing and production decisions.

We model uninsurable business losses within a simple price-setting newsvendor environment.[1] In a world where firms set prices and produce outputs before the realization of their market demand, they will take account of the full distribution of market outcomes contingent to their pricing and production decisions. Each of the contingent outcomes bears either of two loss consequences: excess supply *versus* excess demand. Firms take the costliness of these loss consequences into their pricing and production. We present two complementary motives which firms would act on: cost-compensating motive *versus* loss-balancing motive.

Cost-compensating motive means to reflect the expected cost of business losses likely to incur when producing more. This motive leads firms to add the shadow cost of production to the traditional marginal cost. One important implication is that, even in competitive markets, the additional cost forms wedges between prices and the traditional marginal costs. Consequently, the price–quantity relation showing a firm's willingness to supply comprises the additional cost schedule ridden over the traditional supply curve. In this paper, to avoid confusion, we refer this extended supply curve to ''*offer curve*.'' It slopes upward in the conventional price–quantity coordinate, reflecting the principle that the higher cost a firm must bear, the higher price it wants to charge. Let us call this generalized principle of marginal cost pricing ''effective cost channel,'' or, in short, ''*cost channel*.'' Indeed, this channel of cost compensation has been known in the various contexts: for example, Prescott (1975), Carlton (1979), Eden (1990), Rotemberg and Summers (1990), and Greenwald and Stiglitz (1993).

Loss-balancing motive means to countervail higher business losses likely to incur when producing more. This motive leads firms to lower prices when increasing production. According to the cost channel stated above, firms will associate higher production levels with higher prices to the extent that increased production implies a rise in the odds of excess supply over excess demand. However, higher production levels lower the odds of excess demand at the same time. Profit-maximizing firms will choose production levels to balance the probabilities of being in the excess demand and excess supply states, with the relative costliness of the two states that depend on the price of output sold and the price of output not sold. Increasing the price of output sold makes the excess demand state more appealing relative to the excess supply state, leading the firms to lower production levels. This results in a negative price–quantity relationship. Let us refer it to a ''*hedge curve*,'' and call the generalized principle of loss-balancing pricing ''*hedging channel*.'' This hedging channel would naturally arise where firms access to incomplete spanning of assets over their stochastic goods market outcomes. However,

1. As is named after the metaphorical example of a street-corner newsstand, a typical price-setting newsvendor model assumes that firms choose prices and order stocks before the realization of market demand. See Petruzzi and Dada (1999) for a review of the newsvendor literature.

it has been largely neglected in the literature. Previous works that have some touch on the loss-balancing role of pricing include Greenwald and Stiglitz (1989), Aguirregabiria (1999), and Petruzzi and Dada (1999).

The two complementary channels work together and determine optimal responses of pricing and production to changing economic conditions. Suppose an interest rate cut by the Fed, which affects market discounting factor. All firms will then value higher their unsold products and thus become more tolerant toward possible occurrence of excess supply state. Through the cost channel, this reduction in the expected losses of excess supply makes firms willing to lower price for every given level of output (i.e., the offer curve shifts downward). At the same time, however, the initial balance of probabilities between the excess supply and excess demand breaks down due to the shock. Consequently, through the hedging channel, firms will rebalance them at a new optimum by taking larger probability of excess supply relative to excess demand and thereby charge higher prices for every given level of output (i.e., the hedge curve shifts upward). One can see the two channels react together uniformly pushing up output but counteracting each other in price dimension. As a result, the level of output rises across all sectors whereas the precise direction of price movements can differ sector by sector.

We take the two complementary channels to the analysis of optimal markup behaviors. A key condition determining whether markups rise or fall during booms concerns the relative responsiveness of the two channels to changing economic conditions. In turn, the responsiveness of the channels relies on the way in which firm's fundamentals such as technology and market power are translated into the relative costliness of business environment in each possible state. Intuitively, firms producing perishables and facing elastic market demand tend to be more worried about unsold products and thus take on a relatively smaller probability of excess supply than those firms producing nonperishables and facing inelastic market demand, other things being equal. The former firms will raise prices in response to a shock like the above one because the shock makes them less anxious about excess supply states for given demand distribution. On the contrary, the latter firms tend to be more concerned about lost sales opportunities in excess demand states and thus will lower prices. As a result, with constant marginal costs normalized across sectors, one will see markups procyclical for firms producing perishables and operating in more competitive environment and countercyclical for firms producing nonperishables and operating in more concentrated markets.

Our theory of markups provides a consistent explanation for the mixed evidence of countercyclical and procyclical behaviors at the aggregate and disaggregate levels. Because markups behave differently across sectors, the aggregate behavior conditional to economywide shocks will depend on the entire distribution of sectors. To the extent that countries are different overall in their cost and market structures, markups may also behave differently at the country level. Moreover, the aggregate markup behavior would have somewhat different properties over a long horizon as an economy experiences measurable shifts in its competition structure, for example, due to institutional and legislative changes.

2. Literature Review

In the literature, the two complementary channels have not been brought into a unified framework. Regarding the cost-compensating motive, this paper shares the idea of effective marginal cost pricing with Eden (1990), Rotemberg and Summers (1990), Greenwald and Stiglitz (1993), and Dana (1998, 1999). In this literature, firms facing uncertain demand schedules add up the shadow cost of contingent outcomes to the traditional marginal cost of production. For example, Eden (1990) shows that if we expand the commodity space to include sales probabilities, the optimality condition generalizes the standard marginal cost pricing.

Regarding the loss-balancing role of pricing in goods market, we find three relevant works. In the newsvendor literature of operations research, Petruzzi and Dada (1999) attempt to reconcile different specifications of demand uncertainty by invoking the role of pricing in choosing coefficient of variation of random demand schedules and show the presence of price premium above the certainty benchmark price regardless of demand specifications. Aguirregabiria (1999) also considers the price-setting newsvendor problem with fixed ordering costs, which naturally induce the classical (s, S) inventory rule. In this model, as a theory of clearance sales, the fixed ordering costs increase the stockout risk and make unwanted accumulation of inventories unavoidable and thereby, when combined with additional price adjustment costs, help to explain the coexistence of long periods without nominal price changes and short periods with massive price markdowns. Our work nests such an intrinsic relation between price-setting and stockout risk. Greenwald and Stiglitz (1989) take a more direct approach to the loss-balancing motive in goods markets. To explain real price rigidity, their work employs the notion of "portfolio" within the mean-variance analysis and derives a firm's optimal decision comprising two decision instruments, pricing and production, which differently affect expected returns and risks. Although the present paper shares either modeling device or ideas of managing risk with these previous works, it identifies two distinct complementary channels of pricing and production within a single framework.

Our objective is to explain the behaviors of markups using the two complementary channels of pricing and production. The proposed mechanism is in sharp contrast to the existing literature of markups. Much of the existing explanation resorts solely to the waxing and waning of market power. Within this view, the rate of markup tends to increase in the size of seller's market power. For example, studies based on a game theoretic approach attribute the cyclical markups to the cyclical variations of collusion among competitors (e.g., Green and Porter, 1984; Rotemberg and Saloner, 1986; Rotemberg and Woodford, 1991, 1992). Studies based on switching costs, habit formation, etc., attribute the cyclical markups to the cyclical variations of the elasticity of demand (e.g., Bils, 1989; Klemperer, 1995; Bagwell, 2004; Ravn et al., 2006). Some studies suggest that markups move up and down due to the economies of scale in shopping technologies or in information dissemination that relatively weaken seller' s market power when shopping intensity is high (e.g., Warner and Barsky, 1995; MacDonald, 2000). All these works share common in their reliance on market power subject to cyclical or seasonal variations. In our theory of markups, market power remains limited and indirect in its role for the determination of markup cyclicalities and works only by its implications on the relative costliness of stochastic business outcomes.[2]

2. There are a few more works that illuminate alternative mechanisms of markup pricing other than by cyclical market power. Chevalier and Scharfstein (1996) look into a link between capital market imperfection and goods market pricing. In their work based on customer markets, firms having difficulty in raising external funds during recessions raise markups as a means of helping cash flow while sacrificing their current customer stock. Markup pricing can be also seen as an instrument of marketing in loss-leader models, whereby multi-product firms (or retailers) strategically select high demand items and sacrifice their profit margins at discounted prices to boost customers visits: see, for example, Lal and Matutes (1994) and Chevalier et al. (2003).

3. The Model

3.1. Baseline Setup

Consider a simple price-setting newsvendor environment of Petruzzi and Dada (1999) where a risk-neutral monopolistic firm sets price p and produces output y before the realization of market demand d.[3] Market demand is subject to exogenous random demand factor x as follows:

$$d = xD(p), \tag{1}$$

where price $p > 0$ and x is drawn from the probability distribution function $F(x)$ with a support $[\underline{x}, \overline{x}]$, $0 < \underline{x} < \overline{x} < \infty$, and a finite expected value. $F(\cdot)$ is continuous and differentiable with $f(x) = \frac{\partial F(x)}{\partial x} > 0$ on the interior of the support. $D(\cdot)$ is bounded below by zero and twice-continuously differentiable with $D_p < 0$. Let $\epsilon(p) = -\frac{pD_p(p)}{D(p)}$, the price elasticity of demand. It is assumed that $\epsilon_p(p) = \frac{\partial \epsilon(p)}{\partial p} \geq 0$. This condition implies a standard property that marginal revenue is strictly decreasing.[4] In our model, the condition also serves to limit the exercise of a firm's monopolistic market power (for it states that an increase in prices induces more elastic market demand) and thus underlies the loss-balancing role of pricing and production.

We assume the standard specification of technology. Production is subject to constant returns to scale. Let c denote a constant marginal cost of production and y the level of output. Because pricing and production decisions are made before market demand realizes, firms will find themselves in the state of either excess supply (overproduction) or excess demand (underproduction). Let $n = (y - xD(p))^+$ denote the level of unsold products being defined by the difference between production and realized demand, where $n = y - xD(p)$ for $y > xD(p)$ and $n = 0$ otherwise. Let h denote the economic value of unsold products per unit. In a static setup, it could take on fixed clearance price net other related costs. In a dynamic setup, the economic value of unsold products could be due to the expected present value of inventories, which depends on various factors such as the nature of products (e.g., automobiles vs. salads), the innovation speed of production technology, economy-wide discount rates (e.g., Fed funds rate), and so on. In either setup, the expected value of unsold products is related to a kind of reservation value at which a firm becomes indifferent between selling or retaining sales. For expositional convenience, this paper presents a static model. Indeed, our static model is a reduced-form version which, given our purpose, has an equivalent to a multiperiod model in recursive dynamic environment under certain conditions.

Profits are given by the total sales revenue minus the total cost, which can be expressible as a function of price p, output y, and the realized value of x:

$$\pi(p, y, x) = p\{y - n\} - cy + hn.$$

3. The assumption of risk-neutral firms is not crucial for our results. We adopt it for expositional purpose as it reduces non-linearity of the model. See the Supporting Information for the case of risk-averse firms. The Supporting Information is available from the author's web site and the publisher's.

4. In a certainty model of monopolist with known x, having marginal revenue strictly decreasing in the level of output requires that $\frac{D(p)D_{pp}(p)}{[D_p(p)]^2} \leq 2$. Because $\epsilon_p(p) = p[\frac{D_p(p)}{D(p)}]^2 \left\{\frac{1}{\epsilon(p)} - \frac{D(p)D_{pp}(p)}{[D_p(p)]^2} + 1\right\}$, the condition $\epsilon_p(p) \geq 0$ holds if and only if $\frac{D(p)D_{pp}(p)}{[D_p(p)]^2} \leq \frac{1}{\epsilon(p)} + 1$. Consequently, for $\epsilon(p) > 1$ around the neighborhood of monopolistic pricing and production, the condition implies the standard property that marginal revenue is strictly decreasing.

Then, the monopolist maximizes its expected profit: $\int_{\underline{x}}^{\overline{x}} \pi(p, y, x) dF(x)$.

Let $\hat{x}$ define the lowest admissible value of the realized x needed to clear the market: $\hat{x} = \min[\arg_x\{n = 0\}]$. Because d in (1) is increasing in x, this definition is equivalent to having a value of x such that

$$y = \hat{x} D(p) \tag{2}$$

for given pair of (p, y). Using the threshold value $\hat{x}$, the expected profit is expressible as

$$\begin{aligned}\int_{\underline{x}}^{\overline{x}} \pi(p, y, x) dF(x) &= \int_{\underline{x}}^{\hat{x}} px D(p) dF(x) + \int_{\hat{x}}^{\overline{x}} py dF(x) - cy \\ &\quad + \int_{\underline{x}}^{\hat{x}} h(y - xD(p)) dF(x),\end{aligned} \tag{3}$$

because the sales volume $(y - n)$ contingent to the realization of x is given by $y - n = xD(p)$ for $x \in [\underline{x}, \hat{x})$ and $y - n = y$ for $x \in [\hat{x}, \overline{x}]$.

Assuming the existence of an interior solution to (3), we establish two complementary channels of pricing and production. We start with the optimality conditions for profit maximization.

3.2. The Optimality Conditions

With identity (2), (p, y) and $(p, \hat{x})$ are one-to-one interchangeable (injective mapping) and in particular y is monotone-increasing in $\hat{x}$ for given p.[5] Henceforth, we conduct our analysis over the coordinate $\{(p, \hat{x})\}$. Under the existence assumption of an interior solution, we substitute (2) into (3), reformulate the expected profit as a function of p and $\hat{x}$, and name the function W:

$$\begin{aligned}W(p, \hat{x}) &= \int_{\underline{x}}^{\hat{x}} px D(p) dF(x) + \int_{\hat{x}}^{\overline{x}} p\hat{x} D(p) dF(x) - c\hat{x} D(p) \\ &\quad + \int_{\underline{x}}^{\hat{x}} h\left\{\hat{x} D(p) - xD(p)\right\} dF(x).\end{aligned} \tag{4}$$

The finite integrals above are well defined and differentiable in $(p, \hat{x})$ because D is continuous and differentiable in p and the space for random demand factor x is nonempty, compact-valued, and continuous $(0 < \underline{x} < \overline{x} < \infty)$. Applying the Leibniz rule to this reformulated maximization problem, we obtain the first-order conditions as follows:

$$\begin{cases} W_{\hat{x}}(p, \hat{x}) = \int_{\hat{x}}^{\overline{x}} pD(p) dF(x) - cD(p) + h \int_{\underline{x}}^{\hat{x}} D(p) dF(x) = 0 \\ W_p(p, \hat{x}) = \int_{\underline{x}}^{\hat{x}} \left[xD(p) + pxD_p(p)\right] dF(x) + \int_{\hat{x}}^{\overline{x}} \hat{x}\left\{D(p) + pD_p(p)\right\} dF(x) \\ \qquad + \ h \int_{\underline{x}}^{\hat{x}} \{\hat{x} - x\} D_p(p) dF(x) - c\hat{x} D_p(p) = 0. \end{cases} \tag{5}$$

A sufficient condition for a local maximum to exist is as follows:

5. The threshold $\hat{x}$ will be *implied* by the monopolist's pricing and production decisions. However, *ex ante*, it works like a "point estimator" imparted by the Bayesian decision theory (e.g., the minimum expected loss estimator) and therefore is incorporated into a decision problem.

Lemma 1: *A pair of price and estimate that satisfies $W_{\hat{x}}(p,\hat{x}) = 0$ and $W_p(p,\hat{x}) = 0$ constitutes a local maximum of $W(p,\hat{x})$ under the condition that $\eta(\hat{x})\epsilon(p) > \frac{p}{p-h}$ where $\eta(\hat{x}) = -\frac{\partial\{1-F(\hat{x})\}}{\partial\hat{x}}\frac{\hat{x}}{1-F(\hat{x})} = \frac{\hat{x}f(\hat{x})}{1-F(\hat{x})}$.*

Proof: See Appendix. Essentially, we show the link between the two stated conditions and the second-order sufficient condition for the existence of a local maximum. We will provide an economic interpretation for these conditions after Proposition 2.

Notice first that the condition $W_{\hat{x}}(p,\hat{x}) = 0$ in (5) can be reduced to

$$\underbrace{c}_{\text{MC.Y}} = \underbrace{p\{1-F(\hat{x})\}}_{\text{MR.Y.XD}} + \underbrace{hF(\hat{x})}_{\text{MR.Y.XS}}, \tag{6}$$

because $D > 0$. Also observe that the condition $W_p(p,\hat{x}) = 0$ can be reduced to

$$\underbrace{\hat{x}D(p)\{1-F(\hat{x})\}}_{\text{MR.P.XD}} + \underbrace{\int_{\underline{x}}^{\hat{x}} \left[x\left\{D(p) + (p-h)\,D_p(p)\right\}\right] dF(x)}_{\text{MR.P.XS}} = 0, \tag{7}$$

because $W_p(p,\hat{x}) = \int_{\hat{x}}^{\overline{x}} \hat{x}D(p)dF(x) + \int_{\underline{x}}^{\hat{x}}[x\left\{D(p) + (p-h)D_p(p)\right\}]dF(x) + \frac{D_p(p)}{D(p)}W_{\hat{x}}(p,\hat{x})$. Here, we attach a set of acronyms to help interpret the mathematical terms in the language of economics:

- MC.Y: marginal cost of production,
- MR.Y.XD: expected marginal revenue w.r.t. production decision over excess demand states,
- MR.Y.XS: expected marginal revenue w.r.t. production decision over excess supply states,
- MR.P.XD: expected marginal revenue w.r.t. pricing decision over excess demand states,
- MR.P.XS: expected marginal revenue w.r.t. pricing decision over excess supply states.

MC stands for marginal cost; MR for marginal revenue; Y for w.r.t. production; P for w.r.t. pricing; XD for excess demand; and XS for excess supply. As can be directly read from the acronyms, the characterization of the two optimality conditions are made in the language of marginal analysis.

We present now two complementary channels that underlie the firm's optimal pricing and production decision: cost channel and hedging channel. Our focus is on a connection between the sufficient condition for interior maxima and the economic principles behind the two channels.

3.3. The Two Channels of Pricing and Production

3.3.1. Cost Channel

Equation (6) is the standard equalization condition between the marginal revenue and the marginal cost of production. It states that, for a chosen price, the level of ex ante optimal production will be found where the marginal cost equals the sum of the expected marginal revenues. Again, each of marginal revenue terms is associated with one of the two distinct demand states, XD and XS. Conditional on XD states, one additional unit of production leads to an additional revenue exactly by the chosen price. Ex ante, the economic value of the additional revenue is captured by the term MR.Y.XD. Conditional on XS states, one additional unit of production means one additional accumulation of unsold products, whose economic value is captured by the term MR.Y.XS.

Let us rewrite (6) as follows:

$$p = \left(1 + \underbrace{\frac{F(\hat{x})}{1 - F(\hat{x})}}_{\text{price-cost wedge}} \right) \{c - hF(\hat{x})\}. \tag{8}$$

The price–quantity relation in this new expression is the one that has been proposed by Prescott (1975), Eden (1990), Rotemberg and Summers (1990), and Dana (1998, 1999), who applied the idea of effective marginal cost pricing to the producer decision problem under demand uncertainty. As previously named, c is the traditional marginal cost directly related to production decision (MC.Y), and $hF(\hat{x})$ is the expected marginal revenue w.r.t. production decision over excess supply states (MR.Y.XS), which can be seen as an expected *reservation* value. Obviously, their difference must be positive at an interior optimum: $c > hF(\hat{x})$. At the same time, it holds that $p > c$ at an interior optimum because otherwise the firm would not have produced. So we will see $p > hF(\hat{x})$ at optimum, which is related to a non-speculation condition to rule out speculative motivation such as producing for hoarding not for sales (i.e., to rule out intentional sales refusal).

As a whole, expression (8) demonstrates that the firm prices up over the direct production cost by the *odds* of XS states, $\frac{F(\hat{x})}{1-F(\hat{x})}$. As the firm increases production, the firm has to face higher probability of being in XS states. As the probability of XS increases, the odds also go up. As the odds go up, the firm charges a higher price. Put this mechanism directly in a Bayesian decision context, a firm (a player) doing business (play) by taking on larger odds would compensate for probable loss by increasing the price (prize) that it receives if it sells (succeeds). In short, the marginal cost-revenue equalization condition (8) shows a firm's willingness to supply in the presence of demand uncertainty, whereby the shadow cost of additional production overrides the traditional supply curve. We name condition (8) as an *offer curve* so that one can have the same intuition but avoid confusion with the traditional supply curve.

Proposition 1 (Offer curve): *Given a non-degenerate distribution* $F(\cdot)$, *the set of pairs* $(p, \hat{x})$ *that satisfies condition* (6) *generates an upward-sloping curve over the plane* $\{(p, \hat{x})\}$.

Proof: See Appendix.

A cost-compensating motive is seen in this proposition, which essentially states that the higher the relative costliness of XS a firm faces, the higher the price it wants to mark up. For given economic environment, this motive deduces a positive slope of the offer curve. It also underlies the mechanism that leads the offer curve to react to a changing business environment. For example, a monopolist will lower prices for every given production level when there is a favorable shock to production technology or to reservation value for each given $\hat{x}$: the offer curve shifts downward. We call this motive cost channel.

3.3.2. Hedging Channel

Equation (7) states that, for a chosen level of production, the firm should set a price at which the two expected marginal revenues offset. Conditional on XD states (i.e., conditional on that it sells as many as it produces), a firm could raise the revenue by $y = \hat{x}D(p)$ if it chose price higher. Thus the expected marginal revenue equals the quantity times the probability of XD, as shown in the term MR.P.XD. However, in XS states, it could lose the revenue, whose amount depends on how sensitively demand responds at each possible XS state. Thus the expected value of additional losses across XS states equals the term MR.P.XS as stated. Consequently, the firm's optimal pricing decision presented in equation (7), enunciated in MR.P.XD + MR.P.XS = 0, seeks to balance the relative costliness of XD and XS states by *hedging* losses contingent on one state against losses contingent on the other. Let us name equation (7) as a *hedge curve*.

PROPOSITION 2 (Hedge curve): *Given a non-degenerate distribution $F(\cdot)$, the set of pairs $(p, \hat{x})$ that satisfies condition* (7) *generates a downward-sloping curve over the plane $\{(p, \hat{x})\}$ under the condition that $\eta(\hat{x})\epsilon(p) > \frac{p}{p-h}$.*

PROOF: See Appendix.

For given business environment, a loss-balancing motive leads to a negative slope of the hedge curve. The same motive also applies when the business environment changes and directs the reaction of the hedge curve. For example, when there is a shock that reduces the probability of XS state for every given production level by affecting the reservation value of products, a monopolist who wants to rebalance the probabilities of XS and XD states will take on relatively more XS-probability by increasing output for every given price (or raising prices for every given production level). In brief, the hedge curve shifts outward (or upward) in response to the shock. We call this motive hedging channel.

The given condition in Proposition 2 has been established as the second-order sufficient conditions by Lemma 1. Indeed, it generalizes the standard textbook statement that a monopolist never chooses price in the inelastic range of market demand ($\epsilon(p) > 1$). Notice first that the term $\eta(\hat{x}) = -\frac{\partial\{1-F(\hat{x})\}}{\partial\hat{x}}\frac{\hat{x}}{1-F(\hat{x})}$ is the elasticity of the probability of XD w.r.t. $\hat{x}$: that is, how much the probability of XD decreases in response to an increase in $\hat{x}$.

Now observe that h is the reservation value at the margin. So $p - h$ is a measure of expected marginal loss in revenues conditional on XS states at optimum. And the current price p amounts to marginal loss in revenues conditional on XD states at optimum. Consequently, we can think of the ratio $\frac{p}{p-h}$ as a measure of the relative marginal loss between XD and XS states found at optimum. The relative marginal loss is greater than or equal to one: $\frac{p}{p-h} \geq 1$, where the equality holds if unsold products have no value.

All in all, the given condition states that the firm's optimal pricing-production decision will not be found in the *inelastic* range of the XD probability and market demand. Intuitively, the monopolist will keep producing as long as additional output has little impact on the XD probability and will keep raising price as long as additional price margin has little impact on market demand. We summarize this condition as follows:

$$\underbrace{\eta(\hat{x})}_{\substack{\text{the elasticity of XD probability}\\ \text{w.r.t. production}}} \quad \underbrace{\epsilon(p)}_{\substack{\text{the elasticity of demand}\\ \text{w.r.t. pricing}}} \quad > \quad \underbrace{\frac{p}{p-h}}_{\substack{\text{the relative marginal loss}\\ \text{between XD and XS}}} \tag{9}$$

This is a generalization of the well-known elasticity condition in a certainty model of monopolist.

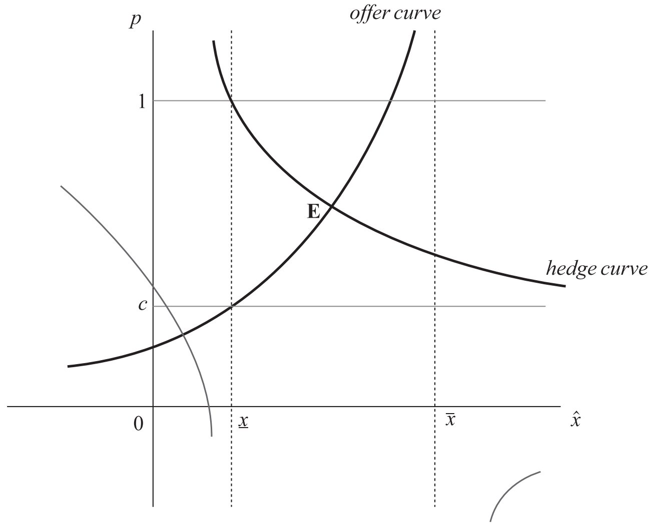


FIGURE 1. THE OFFER AND HEDGE CURVES (EXAMPLE 1)

We finalize this section with an example that illustrates how the hedge and offer curves jointly determine an optimal pair of pricing and production for given business environment.

Example 1: *Let us consider a monopolist that faces market demand* $d = x(1 - p)$, *with* $p \in [0, 1]$ *and* x *uniformly distributed over* $[\underline{x}, \overline{x}]$. *Goods are produced at a constant marginal cost* $c \in (0, 1)$ *and completely perishable yielding* $h = 0$. *Under the assumptions, the offer curve is given by* $p = c\left\{\frac{\overline{x}-\underline{x}}{\overline{x}-\hat{x}}\right\}$ *following from* (6), *and the hedge curve* $p = 1 - \frac{\hat{x}^2-\underline{x}^2}{2(\overline{x}\hat{x}-\underline{x}^2)}$ *from* (7). *As shown in Figure 1, the offer curve is upward-sloping and the hedge curve is downward-sloping. Solving the two simultaneous equations, one can obtain an optimal solution* $(p, \hat{x})$ *where the two curves meet. As indicated by* **E** *in Figure 1, it is unique over the joint space* $[0, 1] \times [\underline{x}, \overline{x}]$.

4. A Map of Markups

4.1. The Business Space: Normalized

By positively relating pricing to production, the offer curve seeks to charge the effective cost of the contingent losses. By negatively relating pricing to production, the hedge curve seeks to balance losses that could arise in XD and XS states. The same principles of cost and hedging channels can be applied to the ways in which the offer and hedge curves react to changing business environment. Differences in technology and market power lead to different responses even to the same changes in business environment. This section studies markup behaviors across different sectors in response to common shocks.

To this aim, one needs to differentiate sector-specific business model from economy-wide business environment. We make it simply by *decomposing* the fundamentals into sector-specific components and economy-wide components. Let $c = \gamma\bar{c}$ and $h = \delta\bar{h}$, where γ and δ are the common components that constitute a part of c and h independent of sector-specific values of $\bar{c}$ and $\bar{h}$. Similarly let $\epsilon(p) = \phi\bar{\epsilon}(p)$, where ϕ is the common part of $\epsilon(p)$ independent of firm-specific function $\bar{\epsilon}(p)$. A set of $(\bar{c}, \bar{h}, \bar{\epsilon})$ can be thought of a business model which differs across sectors; whereas a set of (γ, δ, ϕ) describes a business environment which is common for all sectors. Our comparative statics releases the same shock to business environment and examines its impacts on markups across different business models.

To deal with our analytical results within a two-dimensional plane, we *normalize* the business models over a continuum of sectors as follows: With respect to technology, we use the ratio $\mathsf{a} = \bar{\mathsf{h}}/\bar{\mathsf{c}} \in \mathsf{A}$ to characterize relative "technology" differentiable across sectors. Sectors located near $\mathsf{a} = 0$ would be the ones who produce hardly storable goods and/or implement business models with massive discount of their goods. We use the ratio $\beta = \delta/\gamma$ to characterize the business environment that influences the relative benefits of storage over production at the aggregate level and/or affects market discount factor for all sectors.[6] With respect to market power, we consider demand with a constant sector-specific elasticity $\bar{\epsilon}$ and use its inverse $\mathsf{m} = 1/\bar{\epsilon} \in \mathsf{M}$ (the Lerner index) to characterize relative "market power" differentiable across sectors. Each sector's true market power is m/ϕ, after adjusted by the economy's overall competition ϕ (e.g., changes in antitrust laws and free-trade agreements, etc.). Sectors located near $\mathsf{m} = 1$ tend to have extreme monopoly power. As $\mathsf{m} \to 0$, sectors approach to business models of perfection competition. Now mapping every firm to its business model on the domain of relative technology and market power $\{(\mathsf{a}, \mathsf{m})\}$, the entire world of business is represented by a rectangular space $\mathsf{A} \times \mathsf{M}$. A formal definition is also available for the space of business environment $\{(\beta, \phi)\}$. But just to say for our comparative statics, it is sufficient to have (β, ϕ) disturbed around the neighborhood of $(1, 1)$.

4.2. Anatomy of Markups

Let us map a continuum of sectors onto the space of $\mathsf{A} \times \mathsf{M}$ in which every sector (a, m) is currently at its optimum. Given its business model (a, m) and business environment (β, ϕ), each sector's optimal (gross) markup rate $\mu = p/c$ and optimal estimate $\hat{x}$ satisfying (6) and (7) are related as follows:

$$\mu = \frac{1}{1 - \underbrace{\frac{\mathsf{m}}{\phi}}_{\text{market power}} \underbrace{R(\hat{x})}_{\text{balancing factor}}}, \tag{10}$$

where

$$R(\hat{x}) = \frac{\hat{x}}{T(\hat{x})}\{1 - F(\hat{x})\} + F(\hat{x}) \tag{11}$$

with

$$T(\hat{x}) = \int_{\underline{x}}^{\hat{x}} x dF(x \mid x < \hat{x}) = \frac{\int_{\underline{x}}^{\hat{x}} x dF(x)}{F(\hat{x})}.$$

Equation (10) decomposes optimal markups into two parts: (i) the traditional measure of market power, $\mathsf{m}/\phi = 1/\epsilon$, and (ii) a component attributable to the presence of

6. For example, β rises in response to a fall in Fed funds rate because investors will discount their future cash flow less (i.e., $\delta \uparrow$) and/or firms will be able to access to borrowing facility at lower financial costs (thereby leading to lower production costs; i.e., $\gamma \downarrow$).

demand uncertainty, $R(\hat{x})$. Let us call $R(\hat{x})$ "balancing factor." Notice that for a standard model of monopolist with known $x = \hat{x}$, the balancing factor equals one, $R(\hat{x}) = 1$, because $F(\hat{x}) = 1$. And thus $\mu = \frac{1}{1-\mathsf{m}/\phi}$: That is, for the certainty case, the markup rate is solely a direct reflection of each firm's own market power over the economy's overall competition. For stochastic demand cases, the balancing factor deviates above from 1: $R(\hat{x}) > 1$.

The balancing factor lies at the center of a firm's hedging motive over losses between XD and XS states. Basically, it is a weighted average of $\frac{\hat{x}}{T(\hat{x})}$ and 1 with a weight vector of the XD and XS probabilities. In turn, $\frac{\hat{x}}{T(\hat{x})}$ can be thought of a "conditional sales ratio" between XD and XS states. Notice that it can be deductively written as follows:

$$\frac{\hat{x}}{T(\hat{x})} = \frac{\frac{\int_{\hat{x}}^{\overline{x}} \hat{x} dF(x)}{1-F(\hat{x})}}{\frac{\int_{\underline{x}}^{\hat{x}} x dF(x)}{F(\hat{x})}} = \frac{\int_{\hat{x}}^{\overline{x}} p\hat{x}D(p)dF(x \mid x \geq \hat{x})}{\int_{\underline{x}}^{\hat{x}} pxD(p)dF(x \mid x < \hat{x})}.$$

Apparently, the numerator (denominator) of the last term is the expected sales revenue conditional on the states of XD (XS). A brief digression to finance will help to appreciate the economic meanings behind the ratio. Consider an index that shows the relative size of the expected values of one's investment portfolio between mutually exclusive sets of events (e.g., booms and recessions). An investor would optimally select it by choosing the composition of assets and the total investment size. Similarly, we can think of the conditional sales ratio here as an index of relative revenue between the two mutually exclusive sets of events (XD versus XS). A firm will choose it by combining two decision instruments; pricing and production. Moreover, the firm will choose different conditional sales ratios over times as business environment changes, like an investor would do as the economy moves on.

Having the conditional sales ratio in it, the balancing factor equation (11) shows an optimal way of balancing losses in the XD and XS states. Having the balancing factor in it, the optimal markup rule (10) implies that the balancing factor and the market power jointly determine the behavior of markups in response to changing business environment.

4.3. A Map of Markups

We are ready to carry out comparative statics by releasing shocks to business environment (β, ϕ) and study their impacts and implications on the behavior of markups across different sectors $\{(\mathsf{a}, \mathsf{m})\}$. In this subsection, we will focus on the case of technological environment β. The result from a change in the economy's overall competition ϕ will be examined in the next section where we conduct numerical exercises with respect to various shocks and model specifications.

Suppose a small positive shock that leads β to β' $(\beta < \beta')$; for example, fall in Fed funds rate, advancements of aggregate technology for production and storage, and so on. Such a shock raises each individual firm's expected marginal revenues associated with production in XS states (MR.Y.XS) and reduces the expected losses of holding one additional unit of unsold product in the XS states. Therefore, through the cost channel, firms become willing to lower price for every given $\hat{x}$. In brief, the offer curve shifts downward. One can easily confirm it by equation (6). At the same time, the shock breaks down the initial balance between the XS and XD states, raising each firm's expected marginal revenue in XS states (MR.P.XS) relative to its expected marginal revenue in XD states (MR.P.XD). Consequently, through the hedging channel, firms rebalance them at a new optimum by charging higher price for every given $\hat{x}$. In brief, the hedge curve shifts upward. One can easily confirm it by equation (7). Having the reactions of the two curves together, we find them uniformly pushing up $\hat{x}$ but counteracting each other in p dimension. By implication, the level of output rises across all sectors whereas the precise movement of markups needs a further investigation sector by sector.

Lemma 2: *For any distribution function $F(x)$, which is non-degenerate, continuous on the closed interval $[\underline{x}, \overline{x}]$ with $0 < \underline{x} < \overline{x} < \infty$, and differentiable on the open interval $(\underline{x}, \overline{x})$, the balancing factor $R(\hat{x})$ is non-monotone in $\hat{x}$ and positively (negatively) sloped in the neighborhood of the lower bound $\underline{x}$ (the upper bound $\overline{x}$).*

Proof: See Appendix. Essentially, we show that $-\infty < \lim_{\hat{x}\uparrow\overline{x}} R_{\hat{x}}(\hat{x}) < 0 < \lim_{\hat{x}\downarrow\underline{x}} R_{\hat{x}}(\hat{x}) < \infty$, where $R_{\hat{x}}(\hat{x}) = \frac{dR}{d\hat{x}}$; and then call for the mean value theorem to claim the existence of at least one point $\hat{x}_0 \in (\underline{x}, \overline{x})$ such that $R_{\hat{x}}(\hat{x}_0) = 0$.

This result is highly general and accommodates arbitrary distribution functions for which the balancing factor might have multiple peaks. However, the balancing factor has a single peak for most of the continuous distribution functions used in practice (e.g., Gamma, log-Normal, Weibull, Uniform, etc.). Henceforth, we draw our attention to the case of single-peaked balancing factors based on the family of regular distribution functions. Then, the continuum of sectors will be divided into two groups, depending on whether a member sector's optimal $\hat{x}$ lies below or above a certain value at which a unique peak of the balancing factor occurs. Of course, each sector's optimal $\hat{x}$ depends on its business model in terms of technology and market power (a, m). Intuitively, firms who produce easily perishable products and face highly elastic demand (i.e., relatively small a and small m) tend to choose relatively small $\hat{x}$. Firms characterized by relatively large a and large m tend to choose relatively large $\hat{x}$.

This sectoral division upon the fundamentals is better appreciated by recalling the sufficient condition established by Lemma 1, which requires to hold that $\eta(\hat{x})\epsilon(p) > \frac{p}{p-h}$ at an interior optimum. As previously discussed, this condition generalizes the well-known elasticity condition of a standard monopolist model and states that optimal pricing-production decision will not be found in the *inelastic* range of the XD probability and market demand. It can be now rewritten in terms of the sectoral location (a, m) as follows:

$$\eta(\hat{x}) > \frac{\mathsf{m}}{\phi}\frac{\mu}{\mu - \beta\mathsf{a}}.$$

Obviously, as we travel from sectors with relatively small a and m to those with relatively large a and m (i.e., toward the northeast from somewhere in the southwest on $\mathsf{A} \times \mathsf{M}$), we will see them choose larger and larger $\eta(\hat{x})$ (i.e., more and more elastic XD probability), which in turn implies larger and larger $\hat{x}$ at their own optimum. Recall that this condition has been employed by Proposition 2 to establish the hedging channel of pricing and production. It links now the sectoral fundamentals (a, m) directly to the hedging channel via the non-monotone property of the balancing factor. We show below that the non-monotonicity of the balancing factor has an important implication for the behaviors of markups across sectors.

Proposition 3 (Markup map): *Consider a world of monopolistic markets* $\mathsf{A} \times \mathsf{M}$ *as described above, where every sector* (a, m) *is currently at its optimum. The world* $\mathsf{A} \times \mathsf{M}$ *is partitioned* *according to the property of markups for each sector* (a, m) *in response to a common shock,* β*, which disturbs all the sectors in the neighborhood of their initial optima.*

Proof: First, we already know from Lemma 2 that for any continuous distribution function $F(x)$, there is at least one point $\hat{x}_0 \in (\underline{x}, \overline{x})$ such that $R_{\hat{x}}(\hat{x}_0) = 0$, and $R_{\hat{x}}(\hat{x}) > 0$ ($R_{\hat{x}}(\hat{x}) < 0$) around the neighborhood of the lower bound $\underline{x}$ (the upper bound $\overline{x}$). Second, for those sectors {(a, m)} whose initial optimal $\hat{x}$ are found where R is positively (negatively) sloped, a common shock leading to a higher optimal $\hat{x}$ will result in a higher (lower) value of R at a new optimum. Third, we also know from (10) that a higher (lower) level of R implies a higher (lower) markup rate for every given level of m. Consequently, for the common shock to β, some sectors reduce markups, whereas some raise.

Central to this result is the optimal adjustment of the balancing factor which is non-monotone across sectors in response to changing economic conditions (summarized by changes in β in the present analysis). As previously discussed, an increase in β raises optimal $\hat{x}$ for every sector. The rise of optimal $\hat{x}$ has two effects on the balancing factor. First, it means to rebalance the XS- and XD-probabilities by increasing the XS-probability with a commensurable decrease in the XD-probability. For a small rise in optimal $\hat{x}$, the change in the balancing factor due to rebalancing probabilities is given by $f(\hat{x})\left\{1 - \frac{\hat{x}}{T(\hat{x})}\right\}$. Notice that this direct effect *reduces* the balancing factor because the conditional sales ratio $\frac{\hat{x}}{T(\hat{x})}$ is greater than 1. Second, on the other hand, the conditional sales ratio itself is subject to change in response to changing business environment. This second effect *raises* the balancing factor because the conditional sales ratio is increasing in $\hat{x}$.

The two effects on the balancing factor sum to

$$R_{\hat{x}}(\hat{x}) = \underbrace{f(\hat{x})\left\{1 - \frac{\hat{x}}{T(\hat{x})}\right\}}_{(-)} + \underbrace{\left[\frac{\hat{x}}{T(\hat{x})}\right]' \{1 - F(\hat{x})\}}_{(+)},$$

whose sign will differ across sectors at their optimal path. Lemma 2 implies that for the sectors {(a, m)} who have chosen relatively small initial values for their optimal $\hat{x}$, the second effect is large enough to offset the first effect. As a result, these sectors will raise their balancing factor at optimum in response to a common positive shock to β, and therefore exhibit a procyclical behavior of markups as implied by (10). As we travel toward those sectors whose optimal $\hat{x}$'s are relatively large, we will see their markups countercyclical in response to an increase in β because the second effect diminishes and the slope of the balancing factor turns negative.

We finalize this section with an illustration of Proposition 3.

Example 2: *Assume x follows a log-Normal distribution with mean 1 and standard deviation 0.5. Figure 2 is based on a small shock to* β *in the common business environment initially set at* $(\beta, \phi) = (1, 1)$. *Figure 2a draws the balancing factor* $R(\hat{x})$ *with respect to optimal* $\hat{x}$ *for each given sector* (a, m). *Figure 2b partitions the space of technology* a *and market power* m *into two areas in response to the shock: one area associates a set of fundamentals* {(a, m)} *with the cases in which firms optimally respond to a small increase in* β *by lowering their prices, and the other area associates another set of fundamentals with the cases in which it is optimal to raise prices. The border line collects those sectors* {(a, m)} *whose initial optimum implies the maximal level of* $R(\hat{x})$ *or* $R_{\hat{x}}(\hat{x}) = 0$.

In this illustration, we map a continuum of sectors onto a unit rectangle $\mathsf{A} \times \mathsf{M} = (0, 1)^2$, by which it is assumed that the reservation value of products cannot exceed

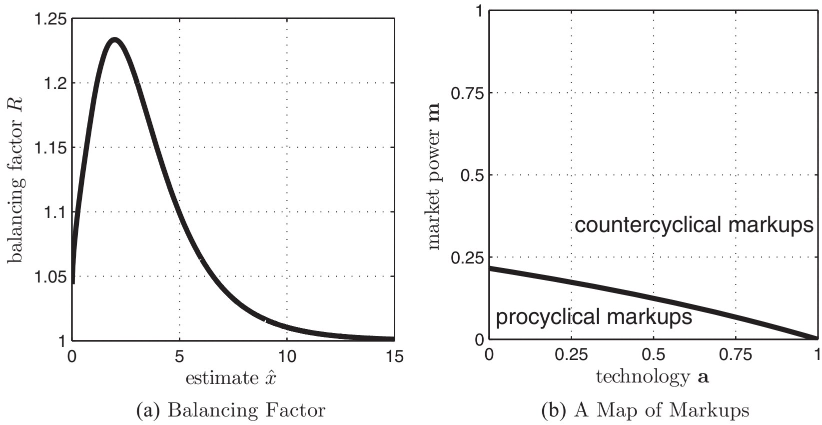


FIGURE 2. CYCLICAL BEHAVIORS OF MARKUPS ACROSS SECTORS (EXAMPLE 2)

Note: *x is assumed to follow a log-Normal distribution with mean 1 and standard deviation 0.5. The figures are based on a small shock to β in the business environment initially set at $(\beta, \phi) = (1, 1)$ for all the sectors on a unit space $A \times M = (0, 1)^2$.*

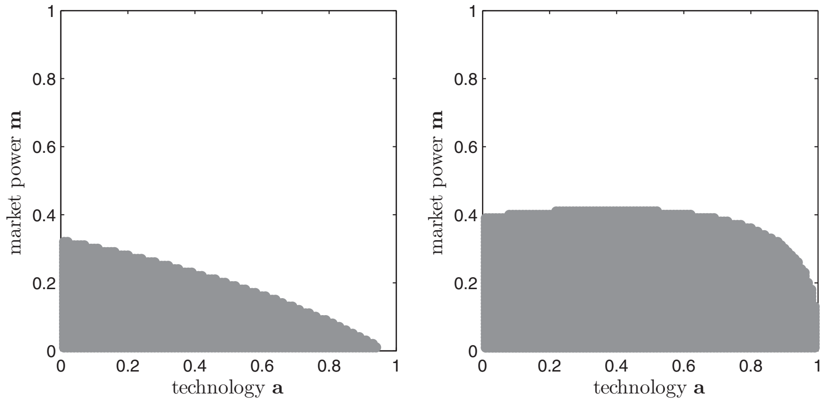
production cost ($0 < \mathsf{a} = \bar{\mathsf{h}}/\bar{\mathsf{c}} < 1$) and the elasticity of demand is greater than one and less than infinity ($0 < \mathsf{m} = 1/\bar{\epsilon} < 1$). However, it is noteworthy that the most popular range of calibration in the literature on the elasticity parameter $\bar{\epsilon}$ lies between 2 and 10 (e.g., 6 in Gali, 2008), which corresponds to the calibration range of market power m between 0.1 and 0.5. Also, note that the precise location of the border line can vary with the initial value of (β, ϕ) and the distribution function F. Nevertheless, the results on the mixed behaviors of markups across sectors essentially remain intact. We discuss further about the robustness of our model and limitations in the next section.

5. Numerical Analysis

5.1. Robustness

This section carries out robustness tests with respect to various shocks and model specifications. We conduct numerical exercises and present the main results by simulation because closed-form solutions are not available for most cases. Another advantage of this approach is that it allows to see the cyclicality of markups directly in terms of output y rather than demand estimate $\hat{x}$. Notice that to some shocks and with some specifications, the optimal response of y may work to the opposite of $\hat{x}$ adjustment if $D(p)$ falls too much in p around a post-shock optimum. It is possible to numerically detect such occasions by recovering every point $(p, \hat{x})$ to (p, y) and to examine the cyclicality of markups according to the joint adjustment of output y and markup μ.

We use a uniform distribution from now on, for the sake of computation efficiency, while keeping the same initial business environment (β, ϕ). A map of markups arises in response to a shock to β, as depicted in Figure 3a. Each gray point indicates that the sector in the corresponding location raises output and markup to a positive shock to β (such as a decrease in market interest rates). And the white area collects the sectors which raise output and lower markups to the same shock. Clearly, the map of markups is essentially identical to the previous map shown in Figure 2b, although the current border line dividing the business world between procyclical and countercyclical lies



(a) Shock to β (b) Shock to ϕ

FIGURE 3. SIMULATED MAPS FROM DIFFERENT SHOCKS

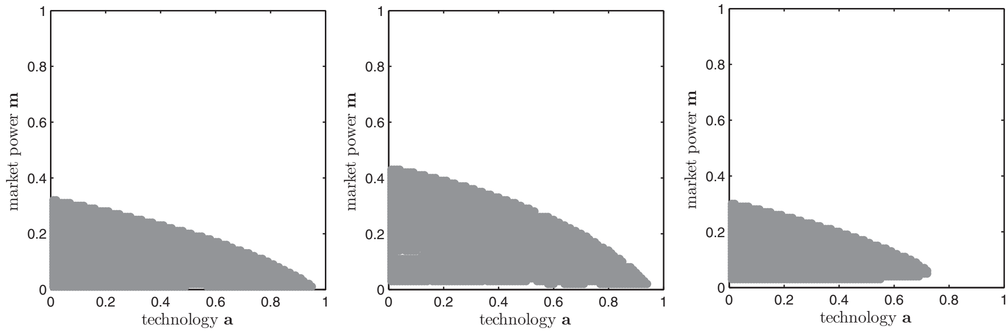
Note: *The figures are based on a uniform distribution with mean 3 and standard deviation 1. The business environment is initially set at* $(\beta, \phi) = (1, 1)$ *for all the sectors. Each shock to* β *and* ϕ *is sized at 0.05 (5% of the initial value). The gray area covers the sectors displaying the procyclical behavior of markups, and the white area showing the countercyclical behavior.*

further northeast relative to the previous one. This minor difference may be attributed to the use of different distribution function, as we control the other conditions being equal.

We continue to explore a map of markups with respect to a shock to the economy's overall competition level ϕ, using the same uniform distribution and initial business environment. We find a similar map, as shown in Figure 3b. The business world is divided into two areas according to the cyclicality of markups. One distinct feature is that the division seems to work tightly between the north and south along the market power dimension m. This may be attributed to the nature of the competition shock. As the competition becomes more fierce, each firm's market power effectively shrinks and all lower their markups. Then, the cyclicality of markups across sectors will be determined by optimal quantity responses. Those firms with relatively weak market powers (i.e., with relatively small m) tend to reduce outputs unless they are competing in markets for easily storable products (i.e., unless having relatively large a). In contrast, the firms who exercise relatively strong market powers in those markets for easily storable products tend to increase output levels. As such, the mechanism behind a border line is essentially identical between shocks to β and to ϕ.

We conduct three more robustness exercises for each case (i) with convex production cost, (ii) for risk-averse firms, and (iii) with explicit stockout cost.[7] We use a quadratic function for total production cost, a logarithmic utility function for firms' risk attitude toward profits, and net stockout cost sized at 1% of unit production cost. Figure 4 presents the results. Again, we find similar maps in both cases. In brief, the two channels and the key mechanism underlying a map of markups that we discover in this paper are present regardless of choice of technology being constant or convex, of decision makers being risk-neutral or risk-averse, and of net stockout cost being free or incurring.

7. Refer to the Supporting Information, for derivation of optimality conditions under different specifications for technology, risk attitude, and explicit cost in excess demand.



(a) With Convex Production Cost (b) For Risk-averse Firms (c) With Stockout Cost

FIGURE 4. SIMULATED MAPS UNDER DIFFERENT SPECIFICATIONS

Note: The figures are based on a uniform distribution with mean 3 and standard deviation 1. The business environment is initially set at $(\beta, \phi) = (1, 1)$ for all the sectors. A shock to β is sized at 0.05 (5% of the initial value). The technology parameter a is set between 0 and 0.95 to mute numerical errors incurred as a approaches to 1. The gray area covers the sectors displaying the procyclical behavior of markups, and the white area showing the countercyclical behavior. Some blank points at the bottom of the map in the figures (b) and (c) are likely due to numerical errors when market power m approaches to zero.

5.2. Limitations

Our theory of markups holds in general under various model specifications with respect to the presence of the two complementary channels. However, the existence result concerning a map of markups conditional on a shock does not hold with all types of shocks and under all model specifications. For example, one may consider shocks to the demand distribution, $F(x)$. Specifically, let us consider shocks to the first two moments of the distribution: (i) a spread-preserving mean shock and (ii) a mean-preserving spread shock. An increase in the first moment can be thought an equivalent rise in demand across sectors; and a decrease in the second moment as a reduction in the degree of demand uncertainty.

We simulate a map by starting with the same uniform distribution and then releasing a shock to each of mean and standard deviation. In response to a positive shock to mean, all firms raise output levels and lower markups. That is, all the sectors display the countercyclical behaviors of markups. In response to a negative shock to standard deviation, which wipes off the density on both sides of the distribution tails, most of the sectors again display the countercyclical markup behaviors. A few exceptions are found in those sectors which have high market power and produce highly storable goods and thus hold pre-shock estimate $\hat{x}$ close enough to the right tail of the initial distribution.[8] We have not tested against all moment shocks and their combinations from the family of all regular distribution functions, but find it safer to make clear that not every type of shocks cause a border across the sectors in the current world of business $\mathsf{A} \times \mathsf{M}$.

Another example for the absence of markup maps is found when the random demand x is additive to demand function $D(p)$.[9] Under the specification, each firm's

8. The reason behind such exceptional complications can be easily understood in terms of first- and second-order stochastic dominance (FSD and SSD). Let F' second-order stochastically dominate F. Some sectors in the rightward of the single-crossing point of F and F' will experience as if being hit by a negative FSD shock when hit by an SSD shock.

9. For derivation of optimality conditions, refer to the Supporting Information.

market power itself is hard to define in a comparable manner, because price elasticity varies with each firm's choice of price p and output y. Keeping it in mind, we assume a linear demand with an additive random factor such that $d = k_0 x + k_1 - k_2 p$. We replace each firm's market power m with the size to slope ratio k_1/k_2. Our numerical analysis shows that a shock to β leads to all procyclical markups across sectors.

Nevertheless, the two limitations above have a common reason. A shock to the demand distribution function causes the balancing factor R function itself to change. And also an additive random demand function yields the balancing factor R as a function of p as well as of $\hat{x}$ and thus makes it float to any changes in the business environment. To either case, one cannot apply the same reasoning developed through Lemma 2 and Proposition 3 in Section 4.3. Therefore, the limitations are clear as a natural consequence of theorization.

6. Empirical Evidence

This section tests the predictions of our theory on the cyclical behavior of sectoral markups over the space of storability technology and market power. We use the U.S. industry data, jointly constructed by the National Bureau of Economic Research (NBER) and U.S. Census Bureau's Center for Economic Studies (CES). The NBER-CES database contains annual industry-level data on output, shipment, inventory, payroll and other input costs, and various industry-specific price indexes. It is released in two versions; one following Standard Industrial Classification (SIC) and the other based on North American Industrial Classification System (NAICS). We use the latter version which covers 473 industries over the period of 1958–2011.[10]

We keep our empirical approach as simple as possible, for the theoretical nature of the present study and given limited space. We estimate proxies for storability and market power from simple regressions and submit them for the analysis of industry-level cyclical properties by correlation. Admittedly, storability and market power proxies may not be free of various estimation issues. In particular, a complete test of our theory requires identification of shocks in respect of nature and sources: For example, one need to identify whether shocks come from supply side or demand side, and whether they work at the industry level or at the aggregate level. We leave such a full-fledged test for future studies.

First, with respect to the industry-level proxy for market power m, we follow Chang et al. (2009) and obtain the price elasticity of industry demand from an IV regression of industry real shipment growth on the rate of industry price change using instruments of TFP growth and the rate of energy price change. These instrumental variables reflect the influence of supply shocks. In theory, demand elasticity is greater than 1 at optimum. In data, elasticity estimates for some industries appear below 1 possibly due to some omitted variables. The empirical counterpart we use for market power is to take an exponential transformation, $\hat{\mathsf{m}}_i = \exp(-\hat{\epsilon}_i)$ for each industry i, so that the market power proxy fits inside [0, 1] for all i's. This makes no harm on the results because the transformation is monotonic. Rather it gives a certain virtue of visual correspondence with our theoretical map.

Second, with respect to the industry-level storability a, there is no straightforward measure available. However, as our model focuses on unexpected inventory holdings,

10. The database and technical notes are available at www.nber.org/nberces/.

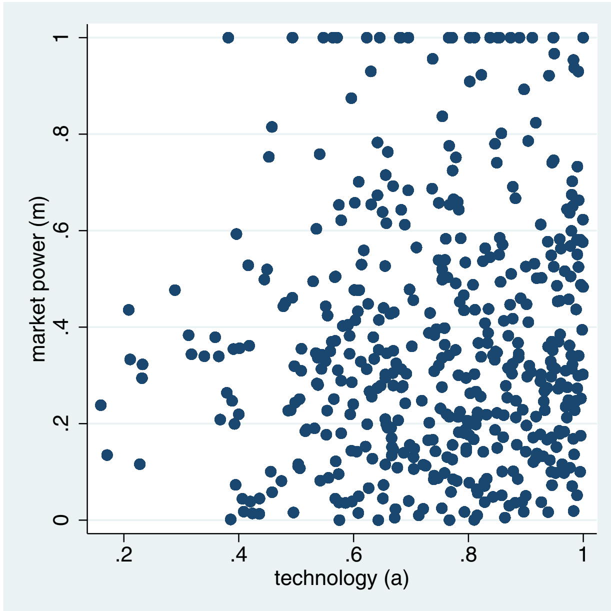


FIGURE 5. LOCATIONS OF U.S. INDUSTRIES

Note: *The figure presents locations of U.S. industries over the space of storability technology and market power.*

it seems reasonable to draw some proxy upon the unexplained part of inventory change from a regression of real inventory growth on real shipment growth. We take "1 minus R-squared" from the regression. Intuitively, unexpected inventory fluctuations will be more endurable for those firms producing more easily storable goods. Therefore, industry R-squared (1 minus R-squared) will show a negative (positive) relationship with latent storability a, because 1 minus R-squared captures the relative volatility of unexplained residuals. By definition, the empirical counterpart of industry-level storability, $\hat{\mathsf{a}}_i = 1 - R_i^2$, fits inside [0, 1] for all i's.

Lastly, with respect to the industry markup measure, we follow the industrial organization literature and use the price-cost margin, that is, the ratio of price minus variable cost to price. Our empirical measure is defined accordingly as (shipment + inventory change - variable cost)/(shipment + inventory change).

Figure 5 shows empirical locations of U.S. industries over the space of storability technology and market power proxies, $\{(\hat{\mathsf{a}}_i, \hat{\mathsf{m}}_i)\}$. Our theory predicts that those industries in the south (north) are more likely to exercise procyclical (countercyclical) markup pricing; so are those in the west (east). Tables I and II provide supporting evidence.

Table I presents the correlation results between the markups and real shipments (in logs) over different range of storability and market power. $Corr|\hat{\mathsf{a}}$ and $Corr|\hat{\mathsf{m}}$ denote contemporaneous correlations between the markups and real shipments for those industries with the storability technology estimated greater than $\bar{\mathsf{a}}$ and with the market power estimated greater than $\overline{\mathsf{m}}$, respectively; that is, correlations conditional on $\hat{\mathsf{a}}_i > \bar{\mathsf{a}}$ and $\hat{\mathsf{m}}_i > \overline{\mathsf{m}}$, respectively. The upper panel traces the correlations along the storability dimension; and the lower panel along the market power dimension.

One who travels toward the east from the west can see more and more industries exhibit countercyclical behavior of markups. The upper panel starts where industries with estimated storability of above 0.55 are doing business. The markups

TABLE I.
RESULTS FROM THE NBER-CES DATA: CORRELATIONS UPON A SINGLE CONDITION

(a) Correlations along the storability dimension

$\overline{a}$	0.55	0.60	0.65	0.70	0.75	0.8	0.85	0.9	0.95
$Corr\|\hat{a}$	0.0335	0.0193	0.0062	−0.0128	−0.0291	−0.0517	−0.0549	−0.0789	−0.2367
Obs.	21,102	19,251	17,486	15,049	13,572	10,901	8,519	6,084	3,546

(b) Correlations along the market power dimension

$\overline{m}$	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.8	0.85
$Corr\|\hat{m}$	0.0026	−0.0392	−0.0640	−0.0667	−0.0390	−0.1272	−0.1935	−0.2087	−0.2507
Obs.	7,753	6,537	5,266	4,321	3,729	2,907	2,489	2,066	1,854

Note: $Corr|\hat{a} = Corr|(\hat{a}_i > \overline{a})$ and $Corr|\hat{m}_i = Corr|(\hat{m}_i > \overline{m})$ denote contemporaneous correlations between the markups and real shipments for those industries with the storability technology estimated greater than $\overline{a}$ and with the market power estimated greater than $\overline{m}$, respectively.

TABLE II.
RESULTS FROM THE NBER-CES DATA: CORRELATIONS UPON JOINT CONDITIONS

	(a)			(b)			(c)		
$\hat{m}_i$	> 0.70	−0.0162	−0.1910	> 0.75	−0.0679	−0.4047	> 0.70	−0.0458	−0.4692
	< 0.70	0.1481	−0.0226	< 0.75	0.1518	−0.0211	< 0.70	0.1594	−0.2008
		< 0.90	> 0.90		< 0.90	> 0.90		< 0.95	> 0.95
		$\hat{a}_i$			$\hat{a}_i$			$\hat{a}_i$	

Note: The table shows contemporaneous correlations between the markups and real shipments for those industries whose estimated storability technology and market power $(\hat{a}_i, \hat{m}_i)$ both satisfy the specified inequality conditions with respect to the pair of thresholds $(\overline{a}, \overline{m})$. For example, in panel (a), 0.1481 is the correlation for those industries with $\hat{m}_i < 0.70$ and at the same time $\hat{a}_i < 0.90$.

look mildly procyclical at start and turn less and less procyclical as we proceed to those with storability above 0.60 and 0.65. While boiling down the membership further, we observe the markups increasingly more countercyclical. However, the countercyclicality does not appear much clear until proceeding to the deep east populated by those industries above 0.95. This is not a downside. It is rather exactly what can be inferred from our theory. In U.S. data, a mass of industries are found in the southeast (see Figure 5). And in our theory, the south tends to exercise procyclical markups.

We now take another journey, south to north, starting where those firms with estimated market power of 0.45 are doing business. As shown in the lower panel, as we proceed further and further toward north, we see more and more industries exercise countercyclical markup pricing: The correlations between markups and real shipments are all negative from the threshold 0.50 on and in absolute terms monotonically increasing (except while shifting to 0.65). This pattern, by and large, also supports our theory of markups.

Table II imposes joint conditions that require estimated storability technology and market power to simultaneously satisfy. The table shows contemporaneous correlations between the markups and real shipments for those industries $\{(\hat{a}_i, \hat{m}_i)\}$ satisfying the specified inequality conditions for the pair of thresholds $(\overline{a}, \overline{m})$. The table has three panels and each panel has 4 domains created by given thresholds. In panel (a), -0.191 from the upper-right domain is the correlation coefficient for those industries with $\hat{m}_i > 0.70$ and $\hat{a}_i > 0.90$. The correlation coefficient from the lower-left domain is 0.1481 for those industries with $\hat{m}_i < 0.70$ and $\hat{a}_i < 0.90$. The results from the two domains are fully consistent with the map of markups theoretically developed and numerically simulated (see Figures 2b, 3, and 4). One can also expect to see acyclical behavior of markups from the rest, provided that the thresholds are imposed to have industries well mixed. The results from panel (a) show indeed so.

We look further into differently paired thresholds. Panel (b) presents the results from raising the threshold of market power to 0.75, while keeping the storability threshold at 0.90 as before. This makes the upper-right domain exhibit much clearer countercyclical behavior, leading to the correlation coefficient over 0.4 in absolute value, without much affecting the rest. Panel (c) goes the other way around: It raises the storability threshold to 0.95 while taking back the threshold of market power as before in panel (a). Again, the upper-right domain shows much clearer countercyclical markups. But, the lower-right one also shows countercyclicality. This is related to the mass distribution in the deep east as depicted in Figure 5. All these facts and patterns consistently support our theory of markups.

7. Conclusions

Firms have access to insurance policies that help them safeguard their business from losses due to various natural disasters but cannot trade a complete set of claims that are contingent on losses incurred through their normal business operation. This paper identifies two complementary channels which firms would act through. One channel is based on cost-compensating motive, by which firms raise prices to reflect higher losses stochastically associated with higher output levels. The other channel is based on loss-balancing motive, by which firms lower prices to countervail higher losses stochastically associated with higher output levels.

We take the two channels to the analysis of markups and show that the behavior of markups differs across sectors depending on the way in which each firm's fundamentals like technology and market power are translated into business environment. The relative responsiveness of the two channels to a shock depends on each firm's fundamental characteristics and leads to a sharp division of markup cyclicality across sectors: It leads to countercyclical adjustment of price-cost markups in sectors producing nonperishable goods with high market power, and procyclical adjustment in sectors having the opposite characteristics of the fundamentals.

This mixed behavior at the disaggregate level suggests that the aggregate markups will depend on the precise distribution of firms on grids of technology and market power. Two implications are immediate: First, if some country comprises more nonfarm business and more concentrated industries than others, the country will be more likely to show countercyclical markups at the aggregate level relative to other countries. Second, if an economy experiences pro-competition institutional changes, the economy will be more likely to show procyclical markups at the aggregate level than before. Although this prediction about cross-sectional and over-time behavioral difference is based on a highly stylized model, the ultimate implication is clear: One will continue to observe mixed behaviors of markups.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Online Supplement

Appendix

Proof of Lemma 1: We will show that under the stated conditions, the second-order total differential d^2W at $(p, \hat{x})$ satisfying the first-order conditions is negative definite: $W_{\hat{x}\hat{x}} < 0$, $W_{pp} < 0$, and $W_{pp}W_{\hat{x}\hat{x}} - W_{p\hat{x}}W_{p\hat{x}}>0$.

Let

$$\begin{cases} L^o(p, \hat{x}) = p\{1 - F(\hat{x})\} - c + \int_{\underline{x}}^{\hat{x}} h dF(x), \\ L^h(p, \hat{x}) = \hat{x}D\{1 - F(\hat{x})\} + \int_{\underline{x}}^{\hat{x}} x\left[D + (p-h)D_p\right] dF(x), \end{cases}$$

where $D = D(p)$ and $D_p = D_p(p)$. In Appendix, we will continue to use these reduced notations to save space. We can then rewrite the first-order conditions as follows:

$$\begin{cases} W_{\hat{x}} = DL^o(p, \hat{x}) = 0, \\ W_p = L^h(p, \hat{x}) + \hat{x}D_pL^o(p, \hat{x}) = 0. \end{cases}$$

And we obtain the second-order partial derivatives:

$$\begin{cases} W_{\hat{x}\hat{x}} = DL^o_{\hat{x}}(p, \hat{x}), \\ W_{\hat{x}p} = D_pL^o(p, \hat{x}) + DL^o_p(p, \hat{x}), \\ W_{p\hat{x}} = L^h_{\hat{x}}(p, \hat{x}) + D_pL^o(p, \hat{x}) + \hat{x}D_pL^o_{\hat{x}}(p, \hat{x}), \\ W_{pp} = L^h_p(p, \hat{x}) + \hat{x}D_{pp}L^o(p, \hat{x}) + \hat{x}D_pL^o_p(p, \hat{x}), \end{cases}$$

where

$$\begin{cases} L^o_{\hat{x}}(p, \hat{x}) = -pf(\hat{x}) + hf(\hat{x}), \\ L^o_p(p, \hat{x}) = 1 - F(\hat{x}), \\ L^h_{\hat{x}}(p, \hat{x}) = D\{1 - F(\hat{x})\} + \hat{x}(p-h)D_pf(\hat{x}), \\ L^h_p(p, \hat{x}) = \hat{x}D_p\{1 - F(\hat{x})\} + \int_{\underline{x}}^{\hat{x}} x\left[2D_p + (p-h)D_{pp}\right] dF(x). \end{cases}$$

It can be easily confirmed that $W_{p\hat{x}} = W_{\hat{x}p}$. Apparently, the signs of $\{L^o_{\hat{x}}, L^o_p, L^h_{\hat{x}}, L^h_p\}$ hold a key to the property of d^2W.

First, make conjecture $\{L^o_{\hat{x}} < 0, L^o_p > 0, L^h_{\hat{x}} < 0, L^h_p < 0\}$. Observe also that $\{L^o(p, \hat{x}) = 0$ and $L^h(p, \hat{x}) = 0\}$ *if and only if* $\{W_{\hat{x}}(p, \hat{x}) = 0$ and $W_p(p, \hat{x}) = 0\}$ because $D > 0$. Accordingly, it is straightforward to confirm $\{W_{\hat{x}\hat{x}} < 0, W_{pp} < 0, W_{pp}W_{\hat{x}\hat{x}} - W_{p\hat{x}}W_{p\hat{x}} > 0\}$ under the conjecture.

Next, we complete our proof by verifying the conjecture. It is immediate that $L^o_{\hat{x}} < 0$ and $L^o_p > 0$. What remains is to show that $L^h_{\hat{x}} < 0$ and $L^h_p < 0$.

(i) $L^h_{\hat{x}} < 0$: Observe

$$\begin{aligned} L^h_{\hat{x}}(p, \hat{x}) &= D\{1 - F(\hat{x})\} + \hat{x}(p-h)D_pf(\hat{x}) \\ &= D\{1 - F(\hat{x})\}\left[1 + \left\{\frac{(p-h)\hat{x}D_p}{\hat{x}D}\frac{\hat{x}f(\hat{x})}{1 - F(\hat{x})}\right\}\right] \\ &= D\{1 - F(\hat{x})\}\left[1 - \frac{(p-h)}{p}\eta(\hat{x})\epsilon(p)\right]. \end{aligned}$$

Because $D\{1 - F(\hat{x})\} > 0$, the stated condition is equivalent to that $L^h_{\hat{x}}(p, \hat{x}) < 0$ on the interior.

(ii) $L_p^h < 0$: Let us rewrite $L_p^h(p, \hat{x})$ using the fact $L^h(p, \hat{x}) = 0$:

$$L_p^h(p, \hat{x}) = -\int_{\underline{x}}^{\hat{x}} \left[\frac{\hat{x} D_p}{\hat{x} D} x \left\{ D + (p-h) D_p \right\} \right] dF(x) + \int_{\underline{x}}^{\hat{x}} \left[x \left\{ 2D_p + (p-h) D_{pp} \right\} \right] dF(x).$$

Using the definition of the price elasticity of demand, we have

$$L_p^h(p, \hat{x}) = \int_{\underline{x}}^{\hat{x}} D_p x \left\{ \frac{(p-h)}{p} \epsilon(p) + 1 + \frac{(p-h)}{p} \frac{p D_{pp}}{D_p} \right\} dF(x).$$

Rearranging it,

$$L_p^h(p, \hat{x}) = \int_{\underline{x}}^{\hat{x}} D_p x \frac{(p-h)}{p} \left\{ \epsilon(p) + 1 + \frac{p D_{pp}}{D_p} \right\} dF(x) + \int_{\underline{x}}^{\hat{x}} D_p x \frac{h}{p} dF(x).$$

The first term is weakly negative because $D_p < 0$ and $\epsilon(p) + 1 + \frac{pD_{pp}}{D_p} \geq 0$ as implied by the assumption $\epsilon_p(p) \geq 0$. The second term is strictly negative because $D_p < 0$. □

Proof of Proposition 1: Let $\left.\frac{dp}{d\hat{x}}\right|_{\text{offer curve (6)}}$ denote the total derivative of p w.r.t. $\hat{x}$ while holding equation (6). It is immediate that

$$\left.\frac{dp}{d\hat{x}}\right|_{\text{offer curve (6)}} = -\frac{L_{\hat{x}}^o(p, \hat{x})}{L_p^o(p, \hat{x})} > 0,$$

from the proof of Lemma 1. Therefore, the offer curve slopes upward. □

Proof of Proposition 2: Let $\left.\frac{dp}{d\hat{x}}\right|_{\text{hedge curve (7)}}$ denote the total derivative of p w.r.t. $\hat{x}$ while holding equation (7). It is immediate that

$$\left.\frac{dp}{d\hat{x}}\right|_{\text{hedge curve (7)}} = -\frac{L_{\hat{x}}^h(p, \hat{x})}{L_p^h(p, \hat{x})} < 0$$

from the proof of Lemma 1. Therefore, the hedge curve slopes downward. □

Proof of Lemma 2: Let $R_{\hat{x}}(\hat{x}) = \frac{dR}{d\hat{x}}$. From (11), we have

$$R_{\hat{x}}(\hat{x}) = f(\hat{x}) - \frac{\hat{x}}{T(\hat{x})} f(\hat{x}) + \left(1 - \frac{T_{\hat{x}}(\hat{x})}{T(\hat{x})} \hat{x}\right) \left\{ \frac{1 - F(\hat{x})}{T(\hat{x})} \right\},$$

where $T_{\hat{x}}(\hat{x}) = \frac{\partial T(\hat{x})}{\partial \hat{x}} = \frac{\{\hat{x} - T(\hat{x})\} f(\hat{x})}{F(\hat{x})}$. Essentially, we will show that $-\infty < \lim_{\hat{x} \uparrow \overline{x}} R_{\hat{x}}(\hat{x}) < 0 < \lim_{\hat{x} \downarrow \underline{x}} R_{\hat{x}}(\hat{x}) < \infty$; and then call for the mean value theorem to claim the existence of at least one point $\hat{x}_0 \in (\underline{x}, \overline{x})$ such that $R_{\hat{x}}(\hat{x}_0) = 0$.

We find that

$$\lim_{\hat{x} \uparrow \overline{x}} R_{\hat{x}}(\hat{x}) = f(\overline{x}) - \frac{\overline{x}}{T(\overline{x})} f(\overline{x}) < 0,$$

because $F(\overline{x}) = 1$ and $0 < \frac{\overline{x}}{T(\overline{x})} < 1$. It is immediate that $R_{\hat{x}}(\overline{x}) > -\infty$ because f and T are finite. We also find that

$$\lim_{\hat{x} \downarrow \underline{x}} R_{\hat{x}}(\underline{x}) = f(\underline{x}) - \frac{\underline{x}}{\lim_{\hat{x} \downarrow \underline{x}} T(\hat{x})} f(\underline{x}) + \left(1 - \frac{\lim_{\hat{x} \downarrow \underline{x}} T_{\hat{x}}(\hat{x})}{\lim_{\hat{x} \downarrow \underline{x}} T(\hat{x})} \underline{x}\right) \left\{ \frac{1}{\lim_{\hat{x} \downarrow \underline{x}} T(\hat{x})} \right\} = \frac{1}{2\underline{x}},$$

because

$$\lim_{\hat{x}\downarrow\underline{x}} T(\hat{x}) = \frac{\lim_{\hat{x}\downarrow\underline{x}} \int_{\underline{x}}^{\hat{x}} x dF(x)}{\lim_{\hat{x}\downarrow\underline{x}} F(\hat{x})} = \frac{\lim_{\hat{x}\downarrow\underline{x}} \frac{\partial \int_{\underline{x}}^{\hat{x}} x dF(x)}{\partial \hat{x}}}{\lim_{\hat{x}\downarrow\underline{x}} \frac{\partial F(\hat{x})}{\partial \hat{x}}} = \underline{x},$$

and

$$\lim_{\hat{x}\downarrow\underline{x}} T_{\hat{x}}(\hat{x}) = \frac{\lim_{\hat{x}\downarrow\underline{x}} \{\hat{x} - T(\hat{x})\} f(\hat{x})}{\lim_{\hat{x}\downarrow\underline{x}} F(\hat{x})} = \frac{\lim_{\hat{x}\downarrow\underline{x}} \frac{\partial \{\hat{x} - T(\hat{x})\} f(\hat{x})}{\partial \hat{x}}}{\lim_{\hat{x}\downarrow\underline{x}} \frac{\partial F(\hat{x})}{\partial \hat{x}}} = 1 - \lim_{\hat{x}\downarrow\underline{x}} T_{\hat{x}}(\hat{x}),$$

where L'Hopital's rule applies to each limit. Therefore, $-\infty < \lim_{\hat{x}\uparrow\overline{x}} R_{\hat{x}}(\hat{x}) < 0 < \lim_{\hat{x}\downarrow\underline{x}} R_{\hat{x}}(\hat{x}) < \infty$.

Finally, the mean value theorem guarantees the existence of at least one point $\hat{x}_0 \in (\underline{x}, \overline{x})$ such that $R_{\hat{x}}(\hat{x}_0) = 0$, because $R_{\hat{x}}(\hat{x})$ is continuous on $(\underline{x}, \bar{x})$. □

References

Aguirregabiria, V., 1999, "The Dynamics of Markups and Inventories in Retailing Firms," *Review of Economic Studies*, 66, 275–308.

Bagwell, K., 2004, "Countercyclical Pricing in Customer Markets," *Economica*, 71, 519–542.

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

Bils, M., 1989, "Pricing in a Customer Market," *Quarterly Journal of Economics*, 104, 699–718.

Blanchard, O., 2009, "The State of Macro," *Annual Review of Economics*, 1, 209–228.

Carlton, D.W., 1979, "Contracts, Price Rigidity, and Market Equilibrium," *Journal of Political Economy*, 87, 1034–1062.

Chang, Y., A. Hornstein, and P.-D. Sarte, 2009, "On the Employment Effects of Productivity Shocks: The Role of Inventories, Demand Elasticity, and Sticky Prices," *Journal of Monetary Economics*, 56, 328–343.

Chevalier, J.A., A.K. Kashyap, and P.E. Rossi, 2003, "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93, 15–37.

———, and D.S. Scharfstein, 1996, "Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence," *American Economic Review*, 86, 703–725.

Dana, J.D., 1998, "Advance-Purchase Discounts and Price Discrimination in Competitive Markets," *Journal of Political Economy*, 106, 395–422.

———, 1999, "Equilibrium Price Dispersion under Demand Uncertainty: The Roles of Costly Capacity and Market Structure," *The RAND Journal of Economics*, 30, 632–660.

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

Eden, B., 1990, "Marginal Cost Pricing When Spot Markets Are Complete," *Journal of Political Economy*, 98, 1293–1306.

Gali, J., 2008, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, Princeton and Oxford.

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li, 2011, "International Prices, Costs, and Markup Differences," *American Economic Review*, 101, 2450–2486.

Green, E.J., and R.H. Porter, 1984, "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87–100.

Greenwald, B.C., and J.E. Stiglitz, 1989, "Toward a Theory of Rigidities," *American Economic Review*, 79, 364–369.

———, and ——— 1993, "Financial Market Imperfections and Business Cycles," *Quarterly Journal of Economics*, 108, 77–114.

Klemperer, P., 1995, "Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62, 515–539.

Lal, R., and C. Matutes, 1994, "Retail Pricing and Advertising Strategies," *Journal of Business*, 67, 345–370.

MacDonald, J.M., 2000, "Demand, Information, and Competition: Why Do Food Prices Fall at Seasonal Demand Peaks?" *Journal of Industrial Economics*, 48, 27–45.

Petruzzi, N.C., and M. Dada, 1999, "Pricing and the Newsvendor Problem: A Review with Extensions," *Operations Research*, 47, 183–194.

Prescott, E.C., 1975, "Efficiency of the Natural Rate," *Journal of Political Economy*, 83, 1229–1236.
Ravn, M., S. Schmitt-Grohe, and M. Uribe, 2006, "Deep Habits," *Review of Economic Studies*, 73, 195–218.
Rotemberg, J., and G. Saloner, 1986, "A Supergame-Theoretic Model of Price Wars during Booms," *American Economic Review*, 76, 390–407.
———, and M. Woodford, 1991, "Markups and the Business Cycle," *Macroecomics Annual*, 6, 63–129.
———, and ———, 1992, "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy*, 100, 1153–1207.
Rotemberg, J.J., and L.H. Summers, 1990, "Inflexible Prices and Procyclical Productivity," *Quarterly Journal of Economics*, 105, 851–874.
Warner, E.J., and R.B. Barsky, 1995, "The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays," *Quarterly Journal of Economics*, 110, 321–352.