Inflation and markups
Theories and evidence from the retail trade sector*

Roland Bénabou

M.I.T., Cambridge, MA 02139, USA and N.B.E.R.

1. Introduction

In this paper I examine how retail markups are affected by high or volatile inflation. While interesting in its own right, this question is best understood in the light of recent work on the real effects of inflation in imperfectly competitive markets. For instance, are inflationary episodes times when consumers' confusion about real prices leads to a rise in monopoly power, or do buyers react to price dispersion and variability by comparing more prices, so that competition intensifies?

The theoretical literature, which can be termed neo-Keynesian due to the central role it ascribes to imperfect competition, is partially reviewed below; Bénabou (1991b) provides a more detailed survey. The main lesson is that once strategic interactions and consumer search are properly taken into account, one cannot take for granted that higher inflation, or even more uncertain inflation, generates welfare losses. Of crucial importance is their effect on market power, which is shown to depend on the size of informational costs. This has an important implication: empirical studies showing a positive relationship between inflation and price dispersion, or even between average inflation and inflation uncertainty, are not sufficient to draw conclusions about the costs of inflation. These variables must themselves be linked to magnitudes which are directly relevant for welfare, and in particular to the extent of competition.

This is what I do here, by constructing a series for the markup in the U.S. retail trade sector and examining how it varies with anticipated inflation, unanticipated inflation, and inflation uncertainty. The markup is derived following the methodology of Rotemberg and Woodford (1991), extended to

take account of intermediate inputs. I focus on retail trade, because it is presumably where search is most relevant.

2. Some theoretical models

Expected inflation. The effects of anticipated inflation have been studied using models where firms face fixed costs of adjusting prices, and operate in an environment of steady inflation. They then optimally follow an (S, s) rule, keeping their nominal price constant while their real price declines from a ceiling S to a floor $s < S$, and then adjusting back to S for a new cycle [Sheshinski and Weiss (1977)]. As inflation rises, S increases and s decreases: price adjustments become larger, and generally more frequent. In equilibrium models, the price dispersion resulting from the staggering of (S, s) rules offers consumers a scope for search, which increases with the rate of inflation [Bénabou (1988, 1992), Diamond (1992)]. This has two opposite effects on welfare. First, when inflation is high more resources will be spent on search; this is a cost of inflation which is frequently mentioned. On the other hand, the increased search pressure implies more competition in the market, pushing real prices S and s down, output and welfare up. In the long run, it also causes some firms to exit, thus reducing the dissipation of rents on fixed costs which characterizes monopolistic competition; on the other hand, having fewer firms in the market may increase the cost or length of search. The overall impact on welfare depends on market structure, and in particular on whether search costs are small (allowing consumers to take full advantage of the widened price range) or large.

Inflation uncertainty. As made clear by the models of Lucas (1973) or Cuckierman (1979), stochastic inflation can generate distortions even in the absence of nominal rigidities, by acting as a source of noise which causes competitive agents to misperceive relative prices. In a more realistic context of imperfect competition and endogenous information, the problem becomes more complex: prices are not Walrasian to start with, and the issue is again whether inflation uncertainty reinforces or weakens monopoly power. This in turn hinges on how it affects the inferences which consumers draw from prices, and the resulting incentives to gather more information [Bénabou and Gertner (1990), Dana (1990), Fishman (1990)]. On the one hand, sellers, being better informed, may use the inflationary noise as a cover behind which to increase their average markups. On the other hand, consumers may react to more noisy prices by gathering more price quotations, forcing down markups. Bénabou and Gertner (1990) formalize these effects in a model where firms are subject to both idiosyncratic and aggregate (inflationary) cost shocks, and where consumers must infer from prices whether or not it is worth searching for a better deal. It shows that where search costs are low,

inflation undertainty can be beneficial by inducing search, but that where they are high it is detrimental to efficiency. Finally, another line of thought linking inflation volatility to decreased efficiency involves repeat purchases: consumers' information about the prices of sellers from whom they purchased previously becomes outdated more rapidly, leaving them less well informed. This allows firms to charge higher prices or to be less efficient [Van Hoomissen (1982), Tommasi (1991)].

Implications. The main conclusions of the two types of models are quite similar: some of the most important welfare effects of inflation and inflation uncertainty involve their impact on market power, through agents' incentives to search. To provide a first empirical assessment of these effects, I shall examine the behavior of markups.

3. Constructing the markup series

Consider a monopolistically competitive industry, composed of firms with production function:

$$Y = \min \{F(H, K) - \Phi, M/\gamma\}, \tag{1}$$

where H is labor hours, K is capital, and M intermediate inputs, which are required in fixed proportion to output. $F(H, K)$ is an increasing, concave function which is homogeneous of degree one, while Φ represents fixed costs generating increasing returns to scale. Cost minimization implies and $M = \gamma Y$ and:

$$pF_H/w = pF_K/r = \lambda, \tag{2}$$

where p is the output price, w the wage, and r the interest rate. The firm's markup μ, defined as the ratio of output price to marginal cost, is therefore:

$$1/\mu = 1/\lambda + s_M, \tag{3}$$

where $s_M \equiv qM/pY = \gamma q/p$ denotes intermediate inputs' share in the value of output and q their price; I denote similarly $s_H \equiv wH/pY$ and $s_K \equiv rK/pY$. Finally, entry leads to the elimination of pure profits: $s_H + s_K + s_M = 1$. Together with (2), this implies:

$$\frac{1}{\mu} = (1 - s_M) \cdot \frac{Y}{Y + \Phi} + s_M, \tag{4}$$

which shows how the fixed costs Φ are reflected in the equilibrium markup.

As shown by Hall (1988), monopoly power is also reflected in the relationship between variations in inputs and outputs:

$$\hat{y} = \mu \cdot (s_H \hat{h} + s_K \hat{k} + s_M \hat{m}),$$ (5)

where hatted variables denote log-deviations from trend, throughout the paper. The difference between the two sides of (5) is the true Solow residual, which for simplicity I take here to be zero. A similar type of accounting holds for value added data. Nominal value added is: $V = wH + rK = pY(1 - s_M)$. Deflating it by the Divisia index ρ, with $\hat{p} \equiv (1 - s_M)\hat{\rho} + s_M \hat{q}$, yields changes in real value added: $\hat{a} = \hat{y} = \mu \cdot (s_H \hat{h} + s_K \hat{k} + s_M \hat{m})$. Hence:

$$\hat{a} = \mu' \cdot (s_H' \hat{h} + s_K' \hat{k}) \quad \text{with} \quad \mu' = \frac{\mu(1 - s_M)}{1 - \mu s_M},$$ (6)

and $s_H' \equiv s_H/(1 - s_M)$, $s_K' \equiv s_K/(1 - s_M) = 1 - s_H'$ denoting labor and capital's shares in value added. Taking markups to be constant over time, Hall (1988) estimates eq. (6) on annual data (1953–1984) for two-digit industries, and obtains sectoral estimates of μ. For the retail sector, on which I focus below, $\bar{\mu}' = 2.355$, corresponding to $\bar{\mu} = 1.403$. Since the issue here is how markups vary with inflation, I shall follow Rotemberg and Woodford (1991) in using Hall's estimate as an average or steady-state value, and constructing from the data a series for deviations of the markup from this level. From (3), we have;

$$\hat{\mu} = \frac{\hat{\lambda} - \lambda s_M \cdot s_M}{1 + \lambda s_M} = \frac{1 - s_M}{1 + (\mu' - 1)s_M} \cdot \left(\hat{\lambda} - \frac{s_M}{1 - s_M} \mu' \cdot s_M \right).$$ (7)

By (1), $s_M = \hat{q} - \hat{p} = (1 - s_M)(\hat{q} - \hat{\rho})$. As to $\hat{\lambda}$, it is obtained by log-linearizing the first-order condition (2), using for F a Cobb–Douglas specification: $F(H, K) = H^\alpha K^{1-\alpha}$, $0 < \alpha < 1$. This leads to $s_H' = \alpha = 1 - s_K'$, and $\hat{\lambda} + \hat{w} - \hat{p} = s_K'(\hat{k} - \hat{h})$. Finally, using (6) and (7):

$$\hat{\mu} = \frac{1 - s_M}{1 + (\mu' - 1)s_M} \cdot (-s_H' - (\mu' - 1) \cdot (s_H' \hat{h} + s_K' \hat{k} + s_M(\hat{q} - \hat{\rho}))),$$ (8)

where μ', s_H', s_K' and s_M can be replaced by their steady-state values $\bar{\mu}'$, $\bar{s}_H'$, $\bar{s}_K'$ and $\bar{s}_M$. The markup computed by Rotemberg and Woodford corresponds to the case $s_M \equiv 0$. A proper treatment of intermediate inputs, however, requires taking account of variations in s_M [see (7)] and of the fact that the real wage implicit in labor's share $-s_H' \equiv \hat{a} + \hat{\rho} - \hat{h} - \hat{w}$ uses the value added deflator instead of the output price; both corrections are reflected in (8).
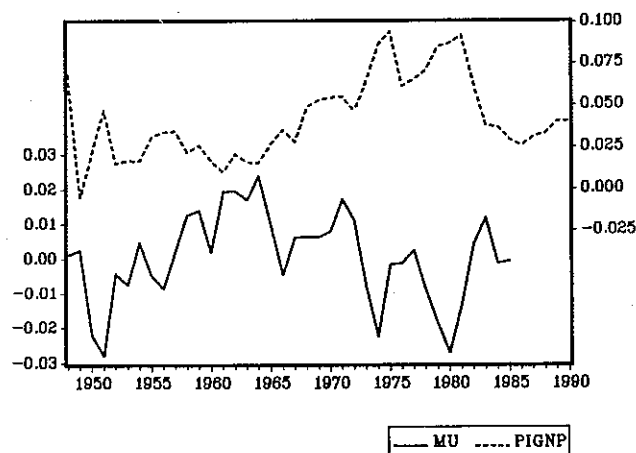
Fig. 1. Inflation and markups.

I use annual data for real and nominal value added, hours, capital, and wages in the retail trade sector, from 1947–1985. It is the same as in Hall (1988), and similar to that used by Rotemberg and Woodford (1991) for other sectors (they do not look at retail trade). For the price q of intermediate inputs I use the producer price index for finished goods.[1] I set $\bar{s}'_H = 1 - \bar{s}'_K$ equal to 0.579, the average share of labor in value added over the sample, and $\bar{s}_M$ to 0.5, which is typical of the literature. Finally, I set $\bar{\mu}'$ equal to Hall's estimate of 2.355, corresponding to an output price exceeding full marginal cost by 40 percent.

4. Sources of variations in markups

Inflation. Figure 1, which plots $\hat{\mu}$ and the rate of inflation in the GNP deflator, conveys a clear message: higher inflation is associated to lower markups. The CPI and PPI give similar results. This picture is confirmed by the regression shown in the first column of table 1, where inflation has a small but significant negative effect: a 10 point rise in inflation reduces the markup by 3.59 percent. Starting from the steady state, it means that price falls from 40.3 percent above marginal cost to 35.3 percent above marginal cost.

---

[1] I also implemented the alternative formulation of (8) which involves $\hat{p}$ and $\hat{q}$ instead of $\hat{p}$ and $\hat{q}$, using for the output price p the deflator of the retail sales series from Citibase; it gave very similar results.

Table 1

Inflation and markups, 1948–1985. Columns (1) to (4) use the GNP deflator, (5) uses the CPI, (6) the PPI.[a]

	(1)	(2)	(3)	(4)	(5)	(6)
cst. $\times 10^3$	6.404	3.858	2.400	-14.37	-0.308	-2.027
	(1.502)	(1.120)	(0.639)	(-1.474)	(-0.050)	(-0.414)
tr $\times 10^3$	0.451	0.369	0.453	0.734	0.557	0.437
	(1.645)	(1.434)	(2.097)	(6.258)	(4.201)	(4.051)
π	-0.359	-0.267				
	(-3.323)	(-2.708)				
$\hat{x}$		0.131	0.143	0.172	0.123	0.154
		(4.768)	(4.280)	(8.012)	(4.969)	(7.538)
$\hat{y}$		-0.333	-0.337	-0.351	-0.306	-0.309
		(-6.596)	(-5.427)	(-7.250)	(-5.400)	(-6.775)
$\bar{\pi}$			-0.329			
			(-2.433)			
$\pi - \bar{\pi}$			-0.241			
			(-2.057)			
$\bar{\sigma}_\pi$			0.169			
			(0.435)			
$E_{-1}(\pi)$				-0.361	-0.249	-0.181
				(-6.340)	(-6.679)	(-7.532)
$\pi - E_{-1}(\pi)$				-0.453	-0.366	-0.202
				(-4.304)	(-6.959)	(-8.585)
$SD_{-1}(\pi)$				1.160	-0.111	-0.038
				(1.660)	(-0.558)	(-0.437)
$\bar{R}^2$	0.286	0.599	0.539	0.727	0.742	0.795
DW	0.957	1.366	1.371	1.758	1.390	1.842

[a]Heteroskedasticity-consistent t-statistics in parentheses.

The business cycle. To properly identify the effect of inflation, one must control for the influence of the business cycle on markups. This is the focus of Rotemberg and Woodford's (1991) study, and I shall use the same variables as them: current industry demand, measured by $\hat{y} = \hat{a}$, and the expected present value of future demand, $\hat{x}_t = E_t(\sum_{j=0}^{\infty} \delta^j \cdot \hat{y}_{t+j+1})$, where a constant discount factor $\delta = 0.9$ is used for simplicity.[2] In oligopolistic markets where firms sustain collusion through trigger strategies, high current demand increases a firm's temptation to deviate, while high future demand increases the collusive profits which will be foregone due to the ensuing price war. Equilibrium markups should thus decrease with $\hat{y}$ and increase with $\hat{x}$. In 'customer markets', market share reacts only slowly to price differentials.

---

[2]I also use the same proxy as them for $\hat{x}$, obtained by projecting sectoral demand on its lagged value and on (detrended) real GNP: $\hat{y}_t = c_1 \cdot \hat{y}_{t-1} + c_2 \cdot \widehat{gnp}_t + v_t$, together with an AR(1) process for real GNP: $\widehat{gnp}_t = c_3 \cdot \widehat{gnp}_{t-1} + \eta_t$, where v_t and η_t are white noise. Hence: $\hat{x}_t = \hat{y}_t/(1 - \delta c_1) + \delta c_2 c_3 \cdot \widehat{gnp}_t/(1 - \delta c_1)(1 - \delta c_3)$.

Thus temporarily high demand leads firms to raise prices without much fear of customer loss, while the expectation of high future demand leads them to compete harder now, in order to attract a clientele which can be profitably exploited later. Hence equilibrium markups should now increase with $\hat{y}$ (unless demand becomes more price-elastic at higher levels) and decrease with $\hat{x}$.

The regression including the business cycle variables is given in the second column of table 1. Inflation's effect is again significantly negative, confirming the previous result. The other notable result is that the cyclical variations of the markup reject the customer market model, and are instead consistent with the implicit collusion model: $\hat{\mu}$ falls with $\hat{y}$ and increases with $\hat{x}$, both effects being statistically significant.

5. Average inflation and inflation variability

Having shown that inflation matters, the next step is to try and understand how. Is the level of inflation relevant in itself, suggesting the presence of nominal rigidities such as those leading to (S, s) rules? Or does inflation matter only through its variability, operating as a source of noise in agents' price information? To try and answer these questions, I split the 1948–1985 sample into eight subperiods of 5 years each,[3] and compute for each of them the average $\bar{\pi}$ and standard deviation $\bar{\sigma}_{\pi}$ of the inflation rate. I then regress $\hat{\mu}$ on these variables, $\pi - \bar{\pi}, \hat{y}, \hat{x}$, a trend and a constant. This procedure is consistent with the idea that the parameters of the inflation process affect market power in steady-state, as in the models of section 2, while business cycle variables and individual inflationary shocks generate deviations from this value. The results, presented in column (3) of table 1, are again clear cut: both average inflation and, to a lesser extent, current deviations from this average value lead to reductions in the markup. Most interestingly, inflation variability has no measurable impact. Finally, the effects of $\hat{y}$ and $\hat{x}$ are unchanged from those of column (2). Using the CPI or PPI instead of the GNP deflator yields identical results.[4]

6. Expected inflation and inflation uncertainty

In this section I make a more sophisticated attempt to identify the effects of uncertainty. I estimate an ARCH forecasting equation [Engle (1982)], generating mutually consistent series for anticipated inflation $E_{-1}(\pi)$, the inflation forecast error $\pi - E_{-1}(\pi)$, and for the latter's standard deviation $SD_{-1}(\pi)$. The specific model used is:

---

[3]Except for the first one, 1948–1950. Using 13 3-year subperiods gave very similar results.

[4]This and all the other regressions in this paper were also run with corrections for serial correlation, as well as in first differences. The results are very similar; see Bénabou (1991b).
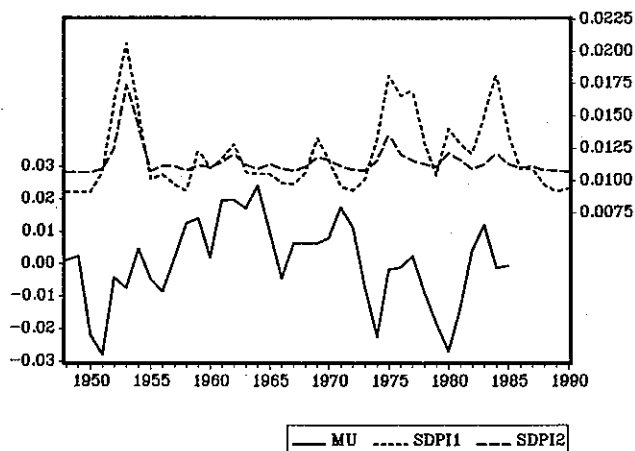
Fig. 2. Inflation uncertainty and markups.

$$\pi_t = Z_t \cdot \beta + \varepsilon_t; \qquad \varepsilon_t \sim N(0, \sigma_t), \qquad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2, \tag{9}$$

where Z_t is a list of variables used to forecast next year's inflation. I use two such lists. The first consists of π_{t-1}, π_{t-2}, a trend and a constant; the second also includes the lagged growth rates of the import deflator and of the hourly wage in manufacturing.[5] Although the graphs are not presented here due to space constraints, plotting $E_{-1}(\pi)$ and $\pi - E_{-1}(\pi)$ clearly reveals that both are negatively correlated with $\hat{\mu}$ [see Bénabou (1991b)]; this is confirmed by the regressions examined below. The relationship between the markup and $SD_{-1}(\pi)$ is illustrated on fig. 2, for both lists of instruments; naturally the smaller information set leads to greater uncertainty. Inflation uncertainty surges up during the early fifties [as in Engle (1983)], at the time of the first oil shock, and in the early eighties during the Volcker deflation; no clear relationship to the markup is apparent. Inflation is measured here using the GNP deflator, but the CPI and PPI give similar results. The regressions shown in columns (4) to (6) of table 1 confirm this picture, and reinforce the results of the previous section: whether we use the GNP deflator, CPI or PPI, both expected inflation and unexpected inflationary shocks depress the markup, with an elasticity of about one-third, and a high level of significance.

[5] Other specifications gave similar results. As in Engle (1983), I ensure that the estimated variance is always positive by constraining the coefficients to decline linearly: $(\alpha_2, \alpha_2) = \beta \cdot (2, 1)$. I estimate α_0 and β by two iterations of generalized least squares [Engle (1982)] rather than by full-information maximum likelihood, for simplicity.

By contrast, inflation uncertainty, measured by the standard deviation of the one-year ahead forecast error, has no measurable impact.

7. Conclusion

Two clear conclusions emerge from this study. First, both expected and unexpected inflation have small but high significantly negative effects on the markups of the U.S. retail sector. Secondly, neither inflation variability nor inflation uncertainty seem to matter. These results are broadly consistent with equilibrium (S, s) models where higher trend inflation, by causing more price dispersion, promotes search and thus intensifies competition. They also suggest that signal-extraction, the obsolescence of price data and other informational considerations do not play any major role. Most importantly, they support the findings of many recent theoretical models that one of the most important welfare effects of inflation is its impact on market power.

References

Bénabou, R., 1988, Search, price setting and inflation, Review of Economic Studies 55, 353–373.
Bénabou, R., 1992, Inflation and efficiency in search markets, Review of Economic Studies, forthcoming.
Bénabou, R., 1991b, Inflation and markups: Theories and evidence from the retail trade sector, MIT Working paper no. 587, Oct.
Bénabou, R. and R. Gertner, 1990, The informativeness of prices: Search with learning and cost uncertainty, MIT Working paper no. 513, Feb.
Cuckierman, A., 1979, The relationship between relative prices and the general price level: A suggested interpretation, American Economic Review 69, June, 444–447.
Diamond, P., 1992, Search, sticky prices, and inflation, Review of Economic Studies, forthcoming.
Engle, R., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987–1007.
Engle, R., 1983, Estimates of the variance of U.S. inflation based upon the ARCH model, Journal of Money, Credit and Banking 15, 287–301.
Fishman, A., 1990, Search, asymmetric information, sticky prices and the Diamond paradox, Mimeo., Dec. (Tel Aviv University).
Hall, R., 1988, The relation between price and marginal cost in U.S. industry, Journal of Political Economy 96, 921–947.
Lucas, R., Jr., 1973, Some international evidence on output–inflation tradeoffs, American Economic Review 63, 326–334.
Rotemberg, J. and M. Woodford, 1991, Markups and the business cycle, MIT Mimeo., May. Forthcoming, NBER Macro Annual.
Sheshinski, E. and Y. Weiss, 1977, Inflation and costs of price adjustment, Review of Economic Studies 44, 287–304.
Tommasi, M., 1991, The consequences of inflation-induced price variability, Mimeo. (University of Chicago).
Van Hoomissen, T., 1988, Price dispersion and inflation: Evidence from Israel, Journal of Political Economy 96, 1303–1314.