DEPARTMENT OF ECONOMICS
DISCUSSION PAPER SERIES

THE L-SHAPED AGGREGATE SUPPLY CURVE AND
THE FUTURE OF MACROECONOMICS

James Forder

Number 486
May 2010

Manor Road Building, Oxford OX1 3UQ
The L-shaped aggregate supply curve and the future of macroeconomics*

James Forder
Balliol College, Oxford
12 May 2010

JEL: B22, E24
Keywords: Phillips curve, wage determination, Keynesianism

I

Almost whenever it is said that early post-War macroeconomics made use of an 'L-shaped aggregate supply curve', the objective seems to be disparagement. The idea is taken to be that below full employment, changes in demand would have no effect on wages, and therefore none on prices, but at full employment, they would affect only prices. Yet soon after the War inflation certainly was a problem, and at levels of unemployment few were inclined to regard as consistent with 'full employment'. So it is no surprise to find that Lipsey (1966) – the second UK edition of his textbook – called the L-shaped approach (p702) 'a simple theory' which was 'commonly used in elementary macro-economics', and said it should be rejected. As he put it in Lipsey (1978), making a point very similar to that of Friedman (1974), these models failed to relate prices and employment in a reasonable way and therefore a further equation was needed to 'close the model'. It was the Phillips curve, of course, that he and Friedman both saw as the next stage in the story, and so superior to the 'L-shape' view did Lipsey find it that he also declared himself unable to understand why 'Keynesian economists at Cambridge' had been so hostile to it.

Particularly in Friedman's hands – notably through Friedman (1977) – the continuation of the story said that the Phillips curve came to be seen as offering a 'menu' to the policymakers and thereby became the basis on which inflationary policy was adopted. As Harcourt (2000) says, it was 'an intellectual and political disaster' that any such trade-off came to be identified with Keynesianism; and indeed – so I have argued in Forder (2010) – it might be added that hardly anyone – Keynesian or otherwise – made the mistakes Friedman alleged to be common.

But these retorts do nothing to breathe plausibility into the L-shaped curve. Conventional economists of the 1950s and 1960s certainly did not believe they faced a stable, exploitable tradeoff, in the manner of the Phillips curve myth, but did they then believe – in the face of the evidence – that inflation could not be a problem below full employment? And if not, what did they believe, before they had the Phillips curve? And what, indeed, did they really believe in that period when, mythology has it, they believed in an exploitable Phillips curve?

* I have benefited from the advice of Ken Mayhew and Margaret Stevens in writing this piece, the comments of Geoff Harcourt, and from the observations of conference participants on a crude and primitive, not to say angular, version of it.
In retrospect one of the notable features of the early post-War discussion of labour economics is the forthrightness, sometimes even the relish, with which it was accepted that wages were, within a certain range, 'indeterminate'. What was meant by this was that the bare economic relations might well set limits to the wage rate that might be paid, but would do not more than that. Reynolds (1949) was typical of the school in thinking that a lower bound was determined by the requirement of attracting sufficient labour, and the upper by considerations of profitability. He did not think it proper to assume strict profit maximization, but in any case, between these limits, he – like many others of the time – thought that economic theory could not determine the wage.

In the case where this indeterminacy arose from the interaction of employer and a trade union, the point might seem to be that of Edgeworth (1881), and some would expect that to be effectively addressed by a development of the theory of bargaining. Certainly the resolution of the indeterminacy of price in bilateral monopoly can be seen as the motivation of a number of early contributions in that area like Pen (1952), and Shackle (1957/1966), or even Zeuthen (1930). But none of them had any pretensions to being more than a beginning, and the labour economists of the 1950s would have treated 'bargaining' as a much more complicated matter, as well as what one might call a genuine bargain – the outcome of an actual human interaction, of parties who knew their interests to be linked, yet opposed; and with human pressures on them. All the wile and guile of give and take; concession and compromise; or entreaty and harangue would be to the fore, as the things which actually resolved the matter. There would be no structural equations of preference, since preferences might even be altered by challenge and conciliation; but in any case, the Bayesian resolution of uncertainty, had it been suggested, would have seemed to offer nothing to the appreciation of bluff and cunning. And indeed, when Nash (1950) had started on the road to a 'theory of bargaining' that would be exemplified in the likes of Binmore and Dasgupta (1987), he himself treated it as highly abstract and lacking in useful applicability.

But the idea of the indeterminacy of the wage was in any case not limited to cases of bilateral monopoly, and an essential in understanding the period is to appreciate the firmness, confidence, and reason with which the marginal productivity theory of wages was rejected even in relation to what might otherwise have been regarded as a competitive market. This point deserves some emphasis since both before and after this period, the marginal product of labour was accorded a prime position in wage determination. In the earlier period, perhaps following particularly from Clark (1899), it became widely accepted as the essential ingredient in wage determination, and received a notable deployment and development in Hicks (1932), chapter 1. And in the 1980s and after it was very much the standard fare of labour economics, and although there might be numerous additional details, leading textbooks like Cahuc and Zylberberg (2004) clearly make marginal product the unchallenged and essential foundation of wage theory.

Yet Lester (1941) ridiculed the theory at length and in Lester (1946a) he expanded the criticism to marginalism more generally, earning swift reposts from Machlup (1946) and Stigler (1947) – the latter in particular being no more gentle with Lester than he
had been with the theory. Hall and Hitch (1939) and Andrews (1949) also questioned the role of marginalism in industrial pricing and the debate that followed is perhaps best remembered – for example, by Lee and Irving-Lessmann (1992) – for that aspect. The best remembered part of Machlup's response came in the mode of what would later be called 'the 'as if' methodology', and denied any value to the results of the kind of questionnaire Lester had used. His position has generally – although not in the view of Lee (1998) – been thought to have had the better of the argument.

Certainly, Lester could be criticised for setting too high a standard for acceptability of a theory – he presented the marginal product theory as if its value depended on the literal truth of a string of extreme assumptions; conceding that its proponents had 'modified' the theory in the light of their falsehood, but still denying it any value. But what is perhaps less appreciated is that in its application to wages, rather than prices, Lester's rejection of the theory had a far broader base than doubts, however severe, about particular assumptions.

There was, in the first place, a mass of evidence that large wage differentials existed in what would otherwise have been regarded as a single labour market. As noted by Reynolds (1946), this had been an observation made by the American War Labor Board when, with a view to arranging for uncontroversial wages, set itself to discover 'the going rate' for various types of labour – they found there was often no such thing. Lester (1946b) reported findings to the effect that on average for comparable jobs in a single locality, the top paying employers paid 50% more than the lowest paying, with a reasonably even distribution of firms in between. He also felt that non-wage compensation would if anything make the discrepancies even starker. Postwar studies reached the same conclusion and Kerr (1950) thought there was 'abundant evidence' on this point, even though he was writing before the publication of Reynolds (1951) which the later literature tended to regard as the locus classicus for this view. Lester (1952) also considered what would later be called 'efficiency wage' interpretations and found they left much to be explained. There were many more of this general type, and Dunlop (1957/1966) was at pains to say that the marginal productivity theory had been widely rejected just because it did not fit the facts.

These views, let it be noted, do not by any means come from peripheral figures – Dunlop, Kerr, Lester and Reynolds were by common consent, and as emphasized in Kaufman (1988), four of the leaders of the field. But one can go beyond this group, and note that even Samuelson (1951) put it pretty candidly. Having gone through a catalogue of wage theories, up to and including marginal productivity theory, and note that even Samuelson (1951) put it pretty candidly. Having gone through a catalogue of wage theories, up to and including marginal productivity theory, rejecting them all, he said that economists simply did not know what determined wages.

What is more, many of the authors concerned sought to test further hypotheses that would clarify the debate. Myers and Maclaurin (1943) p73 found that the limited labour mobility as existed was 'largely ineffective' in equalizing wages. Such equalization as occurred was more a result of union action or minimum wages. In a similar vein, Myers and Shultz (1951) found there was little systematic search for the best opportunity on the part of unemployed workers – they tended to take the first offer they were made. Reynolds and Taft (1956) made a number of points tending to call into question whether differentials even narrowed with time. If the defence of the marginalist propositions was to rest not on their literal truth but on their gravitational
pull, that raised a serious problem. And they too said that it was unionism, rather than the competition resulting from its absence, that tended to move wage structures towards that of the marginalist paradigm, clearly posing a challenge to conventional theory. Many other, often painstaking, empirical studies could be quoted to the effect that in one way or another marginal productivity theory was unable to explain the basic characteristics of wage behaviour.

This is not all. There were also significant theoretical difficulties for the marginal product theory of wages which distinguish it from the argument relating to product prices. One point of perhaps a sentimental kind is the long-standing view that labour is special because it is inseparable from its human provider and the supply and demand for it can hardly avoid being conditioned by people's attitudes. But at the other end of the spectrum, there is a question as to the analytical possibility of the marginal productivity theory being correct – if it is to be taken as determining the payment of all the factors of production. The issue is whether such marginal payments will add up exactly to exhaust the value of the product. The origins of the recognition of this problem are often traced to Wicksteed (1894/1932) but it was notably, if enigmatically, discussed by Robinson (1934). The factor shares will add-up in the appropriate way only under conditions of constant returns and profit maximization – in other words, roughly, only in perfectly competitive equilibrium. It is evident that this problem cannot be addressed along the lines of Machlup since the question is not whether particular behaviour amounts to conformity to a certain theory or not. Rather it is a question of whether it is even arithmetically possible for the postulates of the theory to be correct. It is curious that this point was so rarely noted by the antimarginalists amongst the labour economists. Perhaps their fact-driven analysis would have made the idea of treating all factors in a parallel way an abstraction too far; and they may have been disinclined to give page-room to mathematical considerations even when they supported their arguments. But even so, the point obviously gives strength to their position.

It is also notable that one cannot see, in the relevant period, a tendency for an acceptance of marginalism to grow. Indeed, Pierson (1957), whilst acknowledging the relevance of the theory in the discussion of long-term trends, saw if anything a trend away from its use in practical labour economics, saying

in the area of partial-equilibrium analysis, theorists are putting less emphasis on the rather simple, almost mechanical view of wage determination based on the hypothesis of pure competition and developed in its most elaborate form in the theory of marginal productivity. Increasing attention is being given to formulations based on the noncompetitive hypothesis, such as bilateral monopoly, and on the interplay between competitive and noncompetitive influences.

Even later still, Hicks (1968), commenting on his early work – Hicks (1932) – described what he called 'that terrible first chapter, entitled "Marginal Productivity and the Demand for Labour"' and clearly rejected the theory.
What took the place of any such theory was an attempt to grasp the realities of wage bargaining, even if only to the extent of being able to identity the characteristics of a situation that would be relevant to the outcome. The roots of this approach lay in the work of the American institutionalists and the Webbs (1897/1902), and perhaps in Davidson (1898), and it was firmly based on detailed empirical work in studying the behaviour of actual unions, real firms, and particular bargains, and was certainly not limited to any narrowly defined conception of economics.

For Dunlop (1944) it was 'A fundamental tenet' that

'modes of behavior that are broader than economic theory contribute materially to the understanding of wage determination… the complex ends of rational activity as well as symbolic conduct cannot be overlooked'

Myers (1947) summarized what was needed, saying (p368) that

'In this task the economist may well seek the assistance of his colleagues in psychology, sociology, social anthropology, and political science, because the kinds of questions that need to be answered are not solely, or even largely, economic ones.'

And he suggested a long list of questions about the nature of the labour market, the bargaining institutions and the circumstances and motivations of those involved.

Two related themes emerge strongly in the literature pursuing these questions. One – notably asserted by Slichter (1950) – was that wages would be determined to a significant degree by management policy and discretion. Similarly, Kuhn (1959) suggested that paying high wages was a form of corporate conspicuous consumption, and should be understood in such terms. The other was that the idea of the perceived fairness of the bargain was often taken to be one of genuine causal significance. Like the theory of bargaining, the matter of fairness is something that would later be treated in a rather different manner by, for example, Akerlof and Yellen (1990) although it was also revisited on terms more like the original ones by Solow (1990).

In that earlier literature the point was that perceptions of the fairness of a bargain, although not themselves determinate matters, would be causally effective in wage setting. Fogarty (1961) even thought that actual bargaining brought wages within range of the level implied by the medieval concept of the 'just wage', and Black (1968) thought the 'social and ethical elements in the wage bargain' were then 'well recognised'. Some, indeed, like Wootton (1955), might have given the impression of thinking that ethical postulates could actually over-rule economic law, but most simply accepted that the relevant laws, whatever they might be, formed themselves around the ethical understandings. But they did so, of course, in conjunction with the perceptions and strategies of management. So 'management policy' would be constructed within the economic constraints, but also within the framework of much wider perceptions. Equally, where state-mandated procedures had a much larger role in wage setting, as they did in Australia, it is easy to recognise the importance of fairness in accounts, like that of Russell (1965), of what was done.
Whether the point is taken as being one principally about management policy or as about fairness, it is clear that certain things stood out as tending to play a role in wage determination, and four in particular were the change in the price level; the profitability of the employer; productivity change; and comparison between groups of workers, particularly where there was felt to be a customary link between the wages of the groups.

Sometimes the matter of prices went unremarked, and in any case nothing so mechanical as a mathematical shift of the bargaining relation was contemplated, but particularly as consciousness of ongoing inflation grew, it was clear that increases in the price level were one of the things that had to be taken into account in wage-setting. Ackley (1961) observed as a matter of fact that it seemed to be 'almost automatically accepted' that

‘Wage rates should rise at least as much as the cost of living, quite independently of labor supply and demand’.

Another key consideration was the employer's profit. Reynolds (1949), for example, highlighted this. In the later literature the role of profit was sometimes seen differently Marglin (1984) – gave it a particular role in a more developed theory – and in due course Carlin and Soskice (1990) would treat the point as being that employers have more to lose from a strike when profits are high, and therefore the bargaining position of labour is stronger. That idea was surely always in the background, but the points emphasized by earlier authors like Brown (1962) were much more like those of management policy and related ideas of fairness. When profits were high, it was possible to pay higher wages, and it could well be felt that workers were entitled to share in the good times.

Similarly, improvements in productivity, at least if they were apparent, would be treated as making a case for wage increase. Garbarino (1950) emphasized this, as did Meyers and Bowlby (1958). Again, there could be alternative or mixed understandings of what was the relevant point, but the way of looking at it most consistent with the outlook of time times was that if productivity had increased, it was fair for labour to earn more.

And finally – perhaps the point furthest removed from the modern perceptions in mainstream economics – is the question of the comparison between different groups or workers. Dunlop (1948) pointed to the importance of established patterns of wages as preventing full or easy adjustment to differential changes in productivity. He suggested the concepts of 'wage contours' and 'wage clusters' as ways of theorizing about these things. Ross and Goldner (1950) p254 even made reference to what they called, in this connection, 'the customary equitable criteria of wage adjustment'.

Whilst most of these considerations have their counterpart in the analysis of later years – generally reconstrued as characteristics of efficiency rather than fairness – a striking omission from the list is the question of unemployment. That may be in part because the presumption at the back of the analysts' minds was that policy would not allow large variations in unemployment. But it is also because those things conceived of as being the mechanics of wage bargaining focussed on the circumstances affecting
those conducting the bargain. The general level of unemployment would, on the whole, not do that, although unemployment in particular industries, or of the members of particular unions might. So, Dunlop (1944) made what was at the time an unusually clear-cut suggestion that unions might be seen as seeking to maximize the total wage bill. Even the possibility that they might do this was questioned by Ross (1948) who argued to the effect that unions were in no position to estimate the employer's demand for labour curve so that considerations of the effect of a wage bargain on employment could not be a major factor in their bargaining position.

Ross perhaps initially put his points rather strongly and in a number of places, such as Ross (1950) tried to emphasize that he was not excluding 'economic' factors from consideration. A number of others like Reder (1952) or, later, Kerr (1994), thought that there was nothing to be learned from treating the question as a debate since both sides were making valuable points, without excluding those of the other. But those who took sides tended to side with Dunlop. So, for example, Shultz and Myers (1950) were critical of Ross and maintained that there were plenty of cases where the employment effects of possible wage bargains operated through the sorts of channels Ross felt were important so as to influence a wage bargain. But even they summed up against generalizations, saying, (p380)

'Generalizations which are meant to apply to all unions, therefore, are difficult to support. This is the essential weakness in the position that the employment effect is of no importance in union wage decisions, or that it is found only in the exceptional case. The range of environmental situations in which unions operate at any one time or over a period of time is great enough to permit or impose a variety of pressures and policies, including a consideration of the effect of wage levels on employment in the particular firm or industry.'

And in any case, these are slim pickings by later standards – by the standards of the Phillips curve: Dunlop suggested a rather specific connection of employment in a union to their wage bargaining; Ross doubted it, and on the whole, although not uniformly, the profession sided with Dunlop. Meanwhile, other factors were given much more prominence all round. Employment and unemployment were clearly not taken to be the pre-eminently important variables explaining wage change.

So it is an easy enough matter to give sense to the idea of the 'L-shaped aggregate supply curve'. It is not that, contrary to what Lipsey perhaps suggests, anyone believed it impossible for wages to rise short of full employment, nor even that this was approximately true. Rather, it would be to express the point of view that the issues of wage change and unemployment were separate problems. First of all, there were no very firm laws of the determination of wages – theory could offer some guidance, but evidence showed that within wide limits, there was little to be said about the outcome. And in so far as theory could offer a guide, it did not by any means emphasise unemployment as an important consideration – certainly not the general level of unemployment. There was, then, no 'missing equation', no need to 'close' any model. The model simply did not link the two issues.
There is more to it than that, of course, because the same sort of view of wage bargaining, and indeed of price setting, should be seen as lying behind another idea over which many of the post War economists have been much maligned – the idea of 'cost push' inflation. This idea was, in one way or another, an old one but in the versions of the 1950s and after – despite their great variety – it was construed as arising from the distinctive characteristics of the modern economy.

In one guise the point was that in the context of a policy commitment to full employment, organized labour might seek to raise the wage with inflationary consequences. Robinson (1937/1969) was quickest off the mark, after the appearance of Keynes (1936), with such a warning of the dangers of irresponsible trade unionism. Kalecki (1943) incorporated something of the same idea into the first contribution to what became the 'political business cycle' literature. Beveridge (1944) shared their fears and in order to encourage union responsibility he also proposed a policy of price stability. Fellner (1946) doubted Beveridge's solution but felt that if policy stopped short of actually guaranteeing full employment it would be reasonable to hope that unions and other groups with price-setting power would see their wider interest in the policy. And even Hansen (1947) was not denying the existence of an issue when he took the alternative view that the commitment to full employment would itself induce the labour movement as a whole to be more responsible. Not far removed from these was Slichter (1948), who thought public opinion could be brought to bear to encourage employers to resist wage demands. Viner (1950) – an opponent of the policy – was distinguished from these not by seeing a different problem, or seeing it more clearly, but by not contemplating the possibility of responsible trade unionism.

In slightly later versions, a greater variety of ideas came to be deployed. One branch, exemplified by Slichter (1954) focussed on institutional structure and bargaining arrangements as specific sources of difficulty; another presented by Means (1959) found the problem much more in industrial pricing than union practice. Both of these arguments, of course, were developments of earlier work which pre-dated full employment policy, but the continuity of theoretical perspective is nothing against them. The two were not necessarily separate and Galbraith (1957) certainly saw a complex strategic inter-relation of wage bargaining and price setting.

For the most part, such authors would not have insisted on a completely clear division between cost push and demand pull inflation. Lerner (1959) was unusually definite in placing it at 2% unemployment. Some like Balogh (1958) almost took the view that all inflation was cost-push; but most accepted that, for practical purposes, one often could not be sure how much of a given inflation was due to cost factors and how much to demand factors. Nevertheless, the 'intuitive content' of the distinction – to use the phrase of Fellner (1959) – was clear enough. Demand inflation arose through demand exceeding productive capacity, cost inflation from the raising of prices and wages at lower levels of demand.

One thing that should be clear is that what some authors have taken to be an easy dismissal of the idea of cost inflation, in fact misses the point entirely. Thus, for example, Humphrey (1998), rejected what he called the cost push 'fallacy', saying,
cost-push can at best explain relative prices. It cannot, however, explain the behavior of the aggregate, or general, price level. That is, it cannot do so unless it can show how cost pressures in specific sectors of the economy can markedly influence the money stock, its velocity, or the aggregate level of output—the three variables that jointly determine the general price level. Since there is no reason to think that sectoral cost pressures would materially affect these aggregate magnitudes for any substantial length of time, there is little reason to believe that cost-push theories offer a valid explanation of general price-level movements. Here then is the cost-push fallacy.

One sees how completely this kind of argument misses the point by leaving aside any issues arising in the debate over the Quantity Theory. In the 1950s and 1960s at least, the idea of cost push inflation was, in the first instance, an idea about the problems of economic management in the context of full employment policy. Most definitely, part of the problem was that policymakers were committed to the maintenance of nominal demand sufficient for full employment. This is not a discovery of the critics of the cost push theory, it is of the essence of the perception of its proponents. The issue was whether price stability was compatible with full employment, or whether, rather, unions or business monopolies would make the combination unattainable.

Even in the 1950s, but more notably as time went on, a further issue was raised as to whether the restriction of demand would even be effective in eliminating cost push inflation. With the understanding of wage bargaining which was prevalent, the influence of general unemployment on wage change was slight and numerous authors of the period felt that it would be ineffective. One example is that of Ackley (1958) p624 who said that to avoid cost inflation,

'Aggregate demand must be kept or pushed low enough, and sufficient unemployment created, that unions will not seek or else employers will refuse to grant – strike or no strike – wage increases in excess of productivity increases. This may mean a very considerable body of unemployment.'

The relationship between these positions and the general picture captured by the idea of an 'L-shaped' aggregate supply curve should be clear enough. By making inflation the outcome of the exercise of wage- and price-setting power, the cost push theorists made it an outcome of the choices of the agents concerned, not merely of a mechanical, 'market determined' process. Certainly when the commitment to full employment was part of the setting of the problem, the idea of controlling inflation by reducing demand would be in abeyance. But the general outlook would also take strength from doubts about the effectiveness of such a policy. One would not have to believe – as perhaps Wiles (1973) did – that there is no relation between unemployment and wage change in order to believe that there is no predictable, no stable, or no useful, relation, or that there is no relation of interest to policymakers. And the conclusion would be that, given the will – the will to induce wage bargainers and price setters to co-operate with policy – it would be possible to achieve full employment and stable prices together.

One evident consequence of this is that government attempts to influence wages and prices – through the 'Guideposts' initiated by the Kennedy administration, or attempts at 'incomes policies' in many countries – were not as they may later have seemed,
attempts to prevent prices changing in response to an imbalance of supply and demand. Rather they were attempts to influence the motivation of the bargainers and price setters, and to do so in a manner that acknowledged the importance of perceptions of fairness in wage setting. And belief in their power might have flowed not only from the understanding of wage bargaining, and supportive, if only indicative, econometric evidence, like that in Perry (1966) but also from the experience, as recounted for example by Barber (1975), of President Kennedy managing to induce the steel producers to reverse a price rise they had already announced.

Furthermore, those who continued late into the 1960s or even the 1970s to advocate policy aiming at very high levels of employment should not be presumed to be basing their view on the acceptance of high inflation, or the belief in a stable Phillips curve. When Harrod (1967), for example, said that he preferred the 'Scandinavian objective' of zero unemployment, he was not describing a disregard of inflation or disclosing a failure to appreciate that it would accelerate, but rather presuming that the other proposals he was making would lead to appropriate wage change and there need be no inflation.

None of this, of course, is to suggest that there was uniformity of view. As soon as inflation became a problem there were those who attributed it to excess demand. Selden (1959) was a widely read paper arguing that case in relation to the American economy, and Paish (1958/1966) thought British demand excessive. Phillips (1958) was in effect in this camp. The central point of that paper was that the relation between unemployment and wage change had remained the same over nearly 100 years. It could hardly be, in that case, that there was any reasonable sense in which wages were 'indeterminate' and nor would there by any reasonable argument about the necessity of controlling inflation with demand policy. Phillips' challenge was of course much the greatest – particular inflations always might be due to excess demand, but to say that there was one invariant level of unemployment at which wages were stable challenged the whole understanding of the labour market and its operation and it is easy to see why there would have been widespread aversion to it. It should be clear, I think, in the light of this, why 'Cambridge Keynesians' might be averse to the Phillips curve.

V

The question clearly arises as to how it came about that the true role of the 'L shaped curve', and the real and reasonable role of cost push inflation disappeared from the mainstream of economic understanding. No doubt there are many reasons but one, I would suggest, arises from the way in which discussion of 'the Phillips curve' developed, particularly during the 1960s.

There are two principal strands. One concerns the way the econometric literature developed. It sought, in something of the same manner as Phillips himself, to estimate an equation for the change in average wages. But most early estimations of this kind were in part motivated by the objective of testing the older theories of wage determination. This was never perfectly appropriate since the econometricians dealt with aggregate data, not individual bargains, but nevertheless the idea that, for example, high profits would lead to high wage increases appears to be a testable
proposition, and a number of authors sought to test it. Some of these – like Bhatia (1962) – were thoroughly anti-Phillips in that they denied a role for unemployment in wage determination. But others accepted a role for both unemployment and profits. So, for example, Perry (1964) said

'both because the facts seem to require it and because it seems conceptually preferable, the explanation of wage rate behavior in terms of unemployment rates alone is rejected in favor of a model in which wage changes depend on several variables – profit rates, past changes in living costs, changes in profit rates and the unemployment rate. No attempt is made here to justify formally the hypothesis that these variables are important.'

What would have made it 'conceptually preferable', one might wonder – and yet needing no further explanation – if it were not that these others things were presumed, by the existing, non-econometric literature, to be important?

Several others focused on the idea that some wage bargains were patterned after others – Eckstein and Wilson (1962) was a widely noted example; and Kuh (1967) focused on productivity, declaring it superior to unemployment in predicting wage change. Nearly all of those contributing to what became a large literature studying the econometrics of wage change also included price change as an explanatory variable, and usually did so without much comment.

There were doubters too, of course. Turner (1960) offers one example of the view that the data was inadequate for the kind of studies being undertaken – he questioned whether the unemployment data properly represented the condition of the labour market, and whether wage series adequately captured the variety of forms of compensation and payment that occurred. And Robertson (1962) commented that Phillips' paper pointed to optimism about the possibility of controlling inflation with demand management but that it was difficult to place much reliance on his results because he assumed a 'fixed psychological relation between unemployment and wage change' – just the sort of concern that would flow from an appreciation of the earlier literature.

As time went on, these things changed. The idea that econometrics was an inappropriate tool for studying the question was never likely to be popular amongst those acquiring the skills to perform the estimations, and nor were they likely simply to find no result (or if they did, the work did not make it into print). So econometrics became the accepted way of addressing the issue, at least within economics. Concerns about the quality and appropriateness of data also seem to have faded away.

In the process, contact with the older theories of wage determination was lost. A signal paper in this is Archibald (1969). He described variables that were 'inconsistent with the simple excess demand model' as 'intruders' in the Phillips relation. The terminology, as well as what they took to be the point, was taken up with approval and explained in their widely read and often commended review of the Phillips curve literature by Santomero and Seater (1978). They said,

If anticipated trends are ignored, supply and demand provide an exhaustive list of the factors determining the price of any good. Any factors that shift demand
or supply cause changes in price through changes in excess demand. In a Phillips-type wage equation, the unemployment rate is a proxy for excess demand. Therefore, changes in variables that affect supply or demand should be captured entirely by changes in the unemployment rate, and the other variables themselves should not be entered separately in the wage equation.

Santomero and Seater's ambition is much more to assess the econometrics than it is to evaluate the orientation or theoretical perspective of the authors they consider and their failure to appreciate what the theory of wage determination might have had to offer could hardly be clearer. The misapprehension they were under was not as to which variables capture the effects of supply and demand, but rather the question of whether that is the only role variables might play. They could hardly have written as they did if they had realised that Hicks (1955), by that time thoroughly in tune with the labour economists, had said,

'It has never been the general rule that wage-rates have been determined simply and solely by supply and demand.'

and that in studying the wage bargain,

'it is impossible to make sense of what happens unless one admits that the same general principles of equity in wage-determination are accepted, in a large measure, by both sides.'

By the time of Santomero and Seater – by the time of Archibald, perhaps – contact with this view had, at least in the econometric literature, been lost and it is no wonder that it had become hard to appreciate that an 'L-shaped' curve could be anything but a primitive version of a Phillips curve.

The second strand in the development of thinking arises from the treatment of 'the Phillips curve' in general, non-econometric, discussion. The beginnings of this are to be found in Samuelson and Solow (1960). They are often belittled as having treated the curve as indicating that a persistent inflation could bring a permanent benefit in terms of lower unemployment, and as advocating that policy. They took no such view, in fact, as ought to be apparent to any reader of the paper, if only because of the repeated emphasis they put on the short-run nature of their conclusions. Rather, their paper reflected the point that by the time they wrote it was clear that there was a serious practical difficulty in distinguishing cost push from demand pull inflation. As they noted, it was never realistically possible to say whether a price or a cost increase came first; and even if one took literally the idea that demand inflation could only occur at full employment, it was no great help since there was no unambiguously correct way to determine the numerical value of 'full employment'. Indeed even the conceptual boundary between cost push and demand pull became indistinct. Schultz (1959) attracted a great deal of attention for his suggestion that sectoral decomposition showed that the 1950s inflation was the result of unusually large shifts of demand from some sectors to others, combined with the fact that prices and wages rose more readily in the expanding sectors than they fell in those that were contracting. This 'demand shift' inflation was not quite a true form of cost push inflation but it shared with those arguments the characteristic of arising at less than full employment, and insofar as the problem arose from market imperfections
generating the downward stickiness of wages in contracting sectors it was a 'cost' rather than a 'demand' phenomenon.

So Samuelson and Solow's presentation of what they called a 'modified Phillips curve' depicting a relation between price inflation, rather than wage change, and unemployment, was a reaction to this situation. They used it, not to advocate inflation, but as a way of avoiding arguing about the exact source of inflation – cost push, demand pull, demand shift, or something else – and to characterize the extent of the problem facing American policymakers. What their curve did was graphically represent what they perceived as the incompatibility of price stability and high employment.

It is this view of the Phillips curve – simply as a common sense relation between inflation and unemployment, embodying no particular theory – that informs much of the later discussion of 'the Phillips curve'. It is surely what Modigliani (1977) had in mind when he said that it had

'served to dispose of the rather sterile "cost push" – "demand pull" controversy.'

It is obvious that there is a tension between this and what Archibald and Santomero and Seater felt. If the correct way to estimate a Phillips curve is exclusively through supply and demand variables, it is difficult to see how it can also be a depiction of cost push inflation. But the answer is that there are indeed two kinds of discussion. There is an econometric literature, surveyed by Santomero and Seater, and various other papers in which 'the Phillips curve' is a vaguer relation. Samuelson and Solow made no pretence of presenting econometrics, and 'the Phillips curve' in this line of thinking was an expression of the difficulties of policymaking; it was a recognition that, most probably for reasons like those suggested by Schultze, policy faced deep difficulties in achieving its goals.

But this line of thinking too, led away from the origins of the theories. Samuelson (1961) no doubt contributed by drawing what he called a 'Phillips curve' in a textbook and describing it as showing a cost push problem. This had the effect of suggesting a much more rigid relation than would have been contemplated before. In a way the same is true of those who presented theoretical accounts of cost push inflation as alternative formulations of 'Phillips curves' – Rowthorn (1977) is a late example, but perhaps nevertheless an instructive one. Although he insisted that the cost push considerations could change in magnitude, once he wrote them in an equation it must have seemed that these forces were every bit as 'mechanical' as the neoclassical ones, and contact with the indeterminacy of the earlier literature was lost.

Finally, perhaps it should be noted that there never was an advocate of the L shape theory; there is no good or full statement of that theory from the time when, as I have suggested, the view it describes was widely held. Indeed it is hardly mentioned. Rather, it was a later reconstruction from a period when contact with what that theory was really about had already been lost; a reconstruction which flows from the presumed superiority of the Phillips curve. Indeed, as might have been guessed, although the discussion of the L shaped curve and its inferiority to the Phillips curve
is clear enough in Lipsey (1966), it does not appear at all in Lipsey (1963) – his first edition.

VI

So the 'L shape' curve is a reconstruction from the era of the Phillips curve of the thought of an earlier period, not in itself a feature of the theorizing of that period. But it is, if it is understood properly, a perfectly reasonable reconstruction. It is an expression of the unconnectedness of two problems; or of the fact that the solutions to the problems of unemployment and inflation are different, but themselves more or less unconnected. If one wishes to put it that way, it was a clear denial of there being a 'tradeoff' between inflation and unemployment. This outlook was based on a well-developed conception of the determination of wages, although, of course, not one lending itself to a simple, determinate model; but certainly not, therefore, a simplistic or naïve one.

The recognition of the importance of this kind of view is also an essential in understanding the macroeconomic policy of the period. In this Nelson (2005) is well ahead of others, including some of those most engaged in the study of the history of the Phillips curve. He clearly demonstrates, at least in relation to the 1970s, that an understanding the policy of the time requires an appreciation that an outlook like that of the L-shaped curve explains much of what was done. But something he does not do is show much sign of taking seriously the possibility that these views might have had merit; and he says hardly anything on the question of what would account for their being held.

But the question should be asked as to whether it might not be the theory itself rather than just the reconstruction that is perfectly reasonable. Indeed parts of it are clearly visible in more recent thinking. Although they failed to realise how widespread its recognition had once been, Akerlof and Shiller (2009) regarded the question of fairness in wage setting as a serious omission from the economics of the early twenty-first century; as I argued in Forder (2006), the age of inflation targeting has, despite its genealogy, accepted the reality of cost-push factors in the inflation process; even the thought that the Phillips curve might best be seen as 'horizontal' finds expression in like Kriesler and Lavoie (2005), amongst others; and the broad family of the view of the labour market discussed here can be part of integrated macroeconomic models, as it is in Harcourt (2006).

The step that needs to be taken, however, is something going even beyond these things, namely to recognize that the idea of the Phillips curve itself brings more confusion than enlightenment. So routine has the idea of the Phillips curve itself become that it is enormously difficult to escape the idea that some sort of Phillips curve – a short-run, expectations augmented one, if you like – must be at the center of macroeconomic policy issues. The Phillips curve has become one of those things which 'ramify … into every corner of our minds', although here the problem is in escaping the new ideas so as to be able to accept the old.

But the issue is not – or should not be – whether one feels that in a general way higher levels of demand and employment will be associated with faster inflation. One does not doubt that excessive demand raises the price level; and one does not doubt that
inflation can usually be stopped by a low enough level of demand. It is easy to suppose that one is more or less compelled to accept the existence of a relation in between as well. But this is altogether the wrong way of looking at it. It is the nature of macroeconomics that every variable of interest is ultimately related to every other. Of course a complete specification of the macroeconomy would reveal a relation between inflation and unemployment – no doubt one involving many other variables. But macroeconomic analysis requires that certain relations are identified as the important ones. It is surely the case, for example, that taking very long periods, the average rate of growth and the average rate of unemployment have some relation. There would be no difficulty in devising broadly plausible theories for a variety of different shapes the relation might take. One could even set about testing them. And certainly, a complete specification of the macroeconomy would reveal what that relation is. But we do not tell ourselves that there is therefore a 'missing equation' in that this particular one is not commonly estimated or propounded in textbooks. The question is not whether it is broadly plausible that deeply buried there is some sort of Phillips relation. It is whether there is any evidence or reason to believe, or whether it is anything more than a passing fancy, that there exists such a relation which is definite enough, and stable enough, and vital enough, to be worth anyone's notice.

And there clearly must be much doubt. Despite all the econometrics that has been done, the collective effort to discover the shape or location of the Phillips curve or what makes it move has been none too successful. Griffin (1962) thought that 'Phillips curves' shifted with the actual rate of unemployment. Worswick (1985), using the terminology of the NAIRU said the same thing. So did Galbraith (1997). A decade and more after he thought it 'time to ditch the NAIRU', the only change is the accumulation of evidence pointing in the same direction. And why stop there. Why is it not time to ditch the Phillips curve? If we cannot discover even the rate of unemployment at which inflation is stable, what chance is there of finding the whole Phillips relation?

I suspect that part of the answer is that, since the more thorough-going versions of rational expectations are not to everyone's taste, it can be hard to see how to replace it. If not the Phillips curve, then what? But the older theory points the way. That line of thinking, if it were reconstructed half a century later, could never be identical to its former incarnation. It is, after all, of the essence of the approach that wage change depends on the institutional structure as well as current attitudes, and surely no one will deny that these things are not as they were. But a renewed effort to consider the problems of employment and inflation free of fascination in the Phillips curve would surely bring rewards.

The attitude to the Keynesian era would change as well, and that would certainly be a healthy development. For most purposes, it is true that persuading the economics profession that the Phillips curve has no special place in Keynesianism, or post Keynesianism, or indeed much of a role in serious policy discussion of the 1960s, would be a great advance. The better appreciation of the advantages of the outlook characterised by the L-shaped supply curve might assist in that endeavour. But it could also be that there would be a benefit in changing the post Keynesians' perceptions of the 1960s as well. It was not in any way a period about which excuses need to be made – quite the contrary, in that period lay an approach which fully
reconsidered and reapplied to different circumstances, holds a truer insight, I would suggest, than the Phillips curve ever has.

references


Fellner, W. (1946) Monetary policies and full employment, University of California.

Fellner, W. (1959) Demand inflation, cost inflation, and collective bargaining In The public stake in union power Bradley, P. University of Virginia Press, Charlottesville, VA.


