Effective Demand: Securing the Foundations

Mark HAYES on Olivier ALLAIN (2009) and Jochen HARTWIG (2007)

I very much welcome this opportunity to enter into dialogue with Olivier Allain (OA), whom I have not previously had the pleasure of meeting, and Jochen Hartwig (JH), with whom I have now corresponded for nearly 10 years, I find to my surprise.

Keynes's book is difficult, there is no doubt about that. Even for those of us such as my colleagues, who have spent many hours carefully reading and pondering the text, differences of interpretation remain possible, as 75 years of literature attest. This is not a satisfactory state of affairs, particularly as regards such a fundamental proposition as the principle of effective demand, and I believe we must continue to try and resolve these differences. Nevertheless it is all too tempting for any one of us to claim to have seen the solution and not to listen to the critics who point out the defects in their claim. I myself have come here to listen as well as to offer criticism. Furthermore, my criticism will focus on the extent to which other interpretations conform, in one sense or another, with Keynes and not on their lack of conformity with Hayes.

The central issue is the nature of the equilibrium represented by the point of effective demand, as set out by Keynes in GT Chapter 3. The critical paragraphs are on p. 25:

> Let Z be the aggregate supply price of the output from employing N men, the relationship between Z and N being written $Z = \varphi(N)$, which can be called the *Aggregate Supply Function*. Similarly, let D be the proceeds which entrepreneurs expect to receive from the employment of N men, the relationship between D and N being written $D = f(N)$, which can be called the *Aggregate Demand Function*.
>
> Now if for a given value of N the expected proceeds are greater than the aggregate supply price, *i.e.* if D is greater than Z, there will be an incentive to entrepreneurs to increase employment beyond N and, if necessary, to raise costs by competing with one another for the factors of production, up to the value of N for which Z has become equal to D. Thus the volume of employment is given by the point of intersection between the aggregate demand function and the aggregate supply function; for it is at this point that the entrepreneurs' expectation of profits will be maximised. The value of D at the point of the aggregate demand function, where it is intersected by the aggregate supply function, will be called *the effective demand.*

Let me first try and summarise the common ground. Aggregate and effective demand are about entrepreneurial expectations (OA, 10; JH, 736). Effective demand corresponds to a state of expectation, embodied in a set of expected prices (OA, 10–11; JH, 733). Entrepreneurs operate Marshallian firms under perfect competition, and are concerned with industry and factor prices and with individual convex production functions (OA, 6–7; JH, 730n). There is no radical uncertainty in production at micro-level, so that entrepreneurs can maximise expected profits (OA, 8; JH, 731n). Expectations are binding for a period called by Keynes the 'day' (JH, 729; OA, 2n). There are differences between expectations and outcomes (realised results).

Now the points of difference:

OA claims that Keynes 'assumes that entrepreneurs' short-term expectations are fulfilled' (OA, 3) and that the difficulties in reading GT Chapter 3 stem from Keynes's 'double inconsistency' in reasoning (OA, 4). He argues that Keynes refers to separate aggregate demand and global expenditure functions (OA, 8). Finally, he finds an exception to Keynes's claim that 'the logical theory of the multiplier holds continuously'.

JH argues that the definition of Z should exclude expectation (GT, 24; JH, 730). For JH, aggregate demand alone, not effective demand, embodies the expectation (expected price) (JH, 733). He is silent on the determinants of demand prices, which are semi-exogenous (JH, 737) and insists on the Swedish method (JH, 729).

The text

Somehow I doubt that I will be able to dispose of the argument simply by the analysis of a single sentence! Nevertheless I do think that Keynes's exact words do not support the interpretations put on them by OA and JH. Thus in the passage quoted above, we read

Now if for a given value of N the <u>expected</u> proceeds are greater than the aggregate supply price …

OA reads the passage to mean that equilibrium is reached in practice by a process of convergence of short-term expectations over time, by trial and error. If OA's reading were correct, we would expect to find instead

Now if for a given value of N the <u>realised</u> proceeds are greater than the aggregate supply price …

OA argues that this passage 'leads the reader towards the question of trial-and-error procedure by which entrepreneurs discover where [the point of equilibrium] lies' (OA 4). However a close reading indicates a comparison, not between outcome and expectation, but between expectation and requirement.

The passage clearly describes a procedure for finding equilibrium. However a sequence of causation is not a sequence in time, even though as teachers we often find it necessary to describe the process of reaching static equilibrium step by step, often using a diagram, and students readily confuse this with a dynamic process.

Thus Keynes is describing an instantaneous process of adjustment, in which somehow entrepreneurs collectively choose the amount of employment they wish to offer. OA is too good a scholar of Keynes not to recognise this possibility (OA, 20) and reads in a tacit assumption by Keynes that short-term expectations are fulfilled in order to square the circle. He recognises that Keynes does not make such an assumption explicitly but argues that Keynes is inconsistent in suggesting a trial and error procedure but then supposing it to be unnecessary to specify it (OA, 4).

JH notes that Keynes is referring to a difference between expectation and requirement, rather than expectation and outcome, and argues that clarity requires the removal of the two words 'expectation of' from the definition of aggregate supply price, which reads:

… the aggregate supply price of the output of a given amount of employment is the expectation of proceeds which will just make it worth the while of the entrepreneurs to give that employment (GT, 24)

JH writes 'For Keynes, the supply price is not the market price level an entrepreneur expects, but the proceeds he must have … to satisfy the profit maximising condition' (JH, 730). For JH, it is the demand curve alone which embodies expectations and these are, at least to some extent, exogenous (i.e. not deducible from the level of employment).

The weakness of JH's position is that by removing expectation from the supply side, we lose the possibility of explaining the expectation, i.e. expected price. As he acknowledges, expectations must come from somewhere and if they are not to come from a process of trial and error, as OA suggests, and are to be embodied in the demand curve alone as JH argues, they must be partly exogenous. This is inconsistent

with Keynes's claim to offer a general theory of employment, based on the equilibrium of supply and demand.

The alternative is that Keynes means exactly what he says and that there is no inconsistency. Yet by what mechanism can individual entrepreneurs form a collective and mutually consistent state of expectation? That is the question that has plagued all of us who have thought deeply about it and to which all our papers are addressed. In an interesting sentence, JH writes of the point of effective demand that 'one might conceive of this point as an equilibrium, but it is not some kind of "market equilibrium" ' (JH, 734). The latter point is precisely where we differ.

Equilibrium

A significant difference between OA and JH is that JH follows Chick in holding that the point of effective demand refers to the *ex ante* expectation of proceeds and not to the 'true' equilibrium that may be revealed to be different, *ex post*, if expectations are in error. OA, by contrast, follows Kregel in holding that it is the true (or 'stationary') equilibrium that is the unique point of effective demand.

On this point I support JH to some extent. The implication of OA's analysis is that employment is generally in disequilibrium except in the unlikely event of expectations being fulfilled. In these models, the principle of effective demand does not determine employment at any time (unless we make the tacit assumption that expectations are fulfilled) but only the equilibrium position towards which employment would tend if individual expectations were stable enough to converge. By contrast, Keynes claims to offer a theory of actual employment at any time (*GT*, xxxiii, 4, 245–7) based on the equilibrium of supply and demand (*GT*, xxii–iii, xxxiv–v, 3, 27–30), such that 'today's employment can be correctly described as governed by today's expectations' (*GT*, 50). Yet he himself refers to this as a 'theory of shifting equilibrium' (*GT*, 293).

Here I believe OA (among many others) has been led astray by Kregel (OA, 5n) and I have written a paper on Kregel (1976) which is currently undergoing a contested review process. Although Kregel's construction is elegant and has undoubtedly been persuasive I do not think it has been helpful as a basis for understanding Keynes.

Time and the production period

Different treatments of time are at the root of much of the dispute over the interpretation of *The General Theory*. This in turn relates to the meaning of Keynes's day and period of production (which are explicitly defined, GT, 47n, 287) and the production period (which is not defined in the GT itself, only in drafts), and Keynes's use of the words period, term, short and long. This is the most difficult area in which to persuade others, since it involves looking at the problem in a novel and quite unfamiliar way.

Both JH and OA share a commitment to the Swedish method of *ex ante* and *ex post*, despite Keynes's repudiation of this approach, which JH acknowledges. It is curious that JH then writes that 'this [Swedish] approach is nevertheless essential for the principle of effective demand …' (JH, 735). Although OA mentions only Kregel and not Hicks and the Swedes, their method is the basis of his entire paper.

It is common ground between us that expectations are binding for a period called by Keynes the 'day' (OA, 2; JH, 729), meaning 'the shortest interval after which the firm is free to revise its decision as to how much employment to offer. It is, so to speak, the minimum effective unit of economic time'. I call it Keynes's quantum unit of time.

OA links Keynes's day to his own definition of an 'elementary period … defined by the succession of three operations: hiring precedes production which precedes the sale of output on the market' (OA, 2). Temporary equilibrium is struck in the market (either by price or quantity adjustment) at the end of each elementary period and if expectations have not been fulfilled, hiring and production are adjusted for the next period. 'Several periods are necessary to converge towards a stationary equilibrium' (OA, 5). In similar fashion, JH follows Chick in linking Keynes's day to the production period, of which the essence is that 'it is characterized by the length of time that an entrepreneur is bound by his employment decisions taken at the beginning of that period' and 'plans made at the outset of the period are compared with results realized at the end' (JH, 729). Both OA and JH here adopt, whether consciously or otherwise, the 'week' of Hicks (1939).

JH acknowledges that Keynes attempted (in 1931-32) to create a 'contraption of formulas of process of all sorts of lengths depending on technical factors with income emerging at a given date corresponding to input at an earlier date' (CWXIV, 180), in order to permit a comparison between input and output (expectation and outcome)—and discarded it. Nevertheless JH argues that Keynes's day is an alternative solution to the problem of dividing 'time into periods so that plans can be compared with realized results' (JH, 729).

Yet on p. 287 of *The General Theory* Keynes defines the period of production as having a length n 'if n time-units of notice of changes in the demand for it have to be given if it is to offer its maximum elasticity of employment'. Given Keynes's earlier definition of the day as the time-unit, it cannot be correct to equate the day with a period of production lasting n days. It is clear, both from the GT itself and the later notes and correspondence, that Keynes remains of the view that production processes 'are all of different lengths and overlap with one another' (CW XIV, 185). Consider Chapter 5 and the detailed discussion of processes of different length in section II of Chapter 16. The equation of the day and the production period is tantamount to assuming a uniform production period for all processes.

Yet in what sense is an entrepreneur not bound by his employment decisions at the beginning of a particular process of production? In what sense can the economic distinction be made between the day and the production period? This is much of the burden of GT Chapter 5. Thus Keynes writes

'the original expectations [are not] relevant, which led the firm to acquire the capital equipment and the stock of intermediate products and half-finished materials with which it finds itself at the time when it has to decide the next day's output. Thus, on each and every occasion of such a decision, the decision will be made, with reference indeed to this equipment and stock, but in the light of the *current* expectations of *prospective* costs and sale-proceeds.'

'In the case of short-term expectations … changes in expectation are not, as a rule, sufficiently violent or rapid, when they are for the worse, to cause the abandonment of work on all the productive processes which, in the light of the revised expectation, it was a mistake to have begun; whilst, when they are for the better, some time for preparation must needs elapse before employment can reach the level at which it would have stood if the state of expectation had been revised sooner.'

'Let us consider, first of all, the process of transition to a long-period position due to a change in expectation, which is not confused or interrupted by any further change in expectation. We will first suppose that the change is of such a character that the new long-period employment will be greater than the old. Now, as a rule, it will only be the rate of input which will be much affected at the beginning, that is to say, the volume of work on the earlier stages of new processes of production, whilst the output of consumption-goods and the amount of employment on the later stages of processes which were started before the change will remain much the same as before. In so far as there were stocks of

partly finished goods, this conclusion may be modified; though it is likely to remain true that the initial increase in employment will be modest. As, however, the days pass by, employment will gradually increase.'

'It is evident from the above that the level of employment at any time depends, in a sense, not merely on the existing state of expectation but on the states of expectation which have existed over a certain past period. Nevertheless past expectations, which have not yet worked themselves out, are embodied in the to-day's capital equipment with reference to which the entrepreneur has to make to-day's decisions, and only influence his decisions in so far as they are so embodied.'

Chapter 5 has been neglected for many reasons, not least the insistence on misreading a change in expectations as disappointment in expectations. There has been a further insistence at least since Hansen (1953) on misreading the dynamics of Chapter 5 in terms of the convergence of expectations instead of the adjustment of a heterogeneous capital stock to a new state of expectation. This is despite the text and the accepted Marshallian usage of 'long-period' to refer to capital adjustment.

The global expenditure function

As fuel for his application of the Swedish method, OA draws on Casarosa (1981) and others to make a distinction between the aggregate demand (D) and global expenditure (E) functions. This is motivated, quite properly, by the puzzle as to how the individual expectations of entrepreneurs can lead to a unique macroeconomic state of expectation. OA constructs the D function with reference only to the information available to the individual entrepreneurs, which includes only the expected price and the firm's Marshallian production function. This allows him to construct a curve of demand proceeds (in D, N space) without reference to the propensity to consume, which is concave simply because of diminishing returns. The intersection of concave D and convex Z defines the effective demand for the output of the individual firm. However, *contra* Keynes, this point of effective demand is not unique, since it depends on the state of expectation. There are for OA as many points of effective demand as there are individual expectations of price. Hence the motivation to find a solution for the state of expectation.

At this point OA makes some dubious claims about the text. He states:

'When [Keynes] introduces [the notions of propensity to consume and inducement to invest] in Section II of Chapter 3, he does not write that their sum corresponds to the D function. These two notions are thus of no use in building the aggregate demand function' (OA, 8–9).

'In section II of Chapter 3, Keynes elaborates another function without naming it: the global expenditure function. This one is based on the concepts of propensity to consume and inducement to invest' (OA, 10).

'Keynes defines effective demand as the intersection between the aggregate supply and demand functions. This definition raises a double difficulty: on the one hand there are as many intersection points as states of expectations; on the other hand, the intersection between Z and D does not take into account the behaviour of consumers and investors' (OA, 11).

On the contrary, Keynes writes (GT, section II, Chapter 3, 28–29):

This theory can be summed up in the following propositions:

  (1) In a given situation of technique, resources and costs, income (both money-income and real income) depends on the volume of employment N.

  (2) The relationship between the community's income and what it can be expected to spend on consumption, designated by D_1, will depend on the psychological characteristic of the community, which we shall call its *propensity to consume*. That is to say, consumption will depend on the level of aggregate income and, therefore, on the level of employment N, except when there is some change in the propensity to consume.

(3) The amount of labour N which the entrepreneurs decide to employ depends on the sum (D) of two quantities, namely D_1, the amount which the community is expected to spend on consumption, and D_2, the amount which it is expected to devote to new investment. D is what we have called above the *effective demand.*

(4) Since $D_1 + D_2 = D = \varphi(N)$, where φ is the aggregate supply function, and since, as we have seen in (2) above, D_1 is a function of N, which we may write $\chi(N)$, depending on the propensity to consume, it follows that $\varphi(N) - \chi(N) = D_2$.

(5) Hence the volume of employment in equilibrium depends on (i) the aggregate supply function, φ, (ii) the propensity to consume, χ, and (iii) the volume of investment, D_2. This is the essence of the General Theory of Employment.

I do not see how these two sets of statements can be reconciled. Keynes simply does not make the distinction between D and E attributed to him. OA cannot find a mandate within the text for his solution to the formation of the state of expectation.

What does find support in GT Chapter 5 (but not Chapter 3) is the idea that expectations are based on realised results, which OA refers to as 'conservative expectations' (OA, 2). Keynes writes:

[It] will often be safe to omit express reference to short-term expectation, in view of the fact that in practice the process of revision of short-term expectation is a gradual and continuous one, carried on largely in the light of realised results; so that expected and realised results run into and overlap one another in their influence. … For, although output and employment are determined by the producer's short-term expectations and not by past results, the most recent results usually play a predominant part in determining what these expectations are. … Accordingly it is sensible for producers to base their expectations on the assumption that the most recently realised results will continue, except in so far as there are definite reasons for expecting a change. (*GT*, 50–51)

The causation here runs from realised result to expectation, not the other way. Expectations conform to realised results, not realised results to expectations.

The global expenditure function (E) is the *ex post* aggregate demand of Old Keynesian economics, which generally avoided the problem of expectation formation by assuming fixed prices. It is based on Book III (Chapters 8-10) of *The General Theory* in which Keynes undoubtedly discusses income (an *ex post* concept, discussed at length in GT Chapters 6 & 7) rather than effective demand (an *ex ante* concept). The connection between D and E is made by Keynes, not through sleight of hand in GT Chapter 3, but through the above statements in GT Chapter 5 about the relationship between expected and market prices (expectations and realised results).

The global expenditure function (E) has no place in GT Chapter 3 or in the principle of effective demand. Where OA (incorrectly) states that 'After section I of Chapter 3, Keynes refers to E because he focuses on ensuring coherence on a macroeconomic level' (OA, 12), he should write something like 'After Chapter 5, and for the purposes of Books III and IV, Keynes focuses on income and expenditure rather than effective demand and employment. Effective demand reappears in Book V.'

The multiplier

OA claims (OA, 16) to refute Keynes's statement that 'the logical theory of the multiplier … holds good continuously, without time-lag, at all moments of time' (GT, 122). This conclusion is reached as the by-product of a model of convergence in expectations that occupies nearly half the paper and which OA regards as clarifying a gap in *The General Theory* (OA, 21).

The proposed refutation arises from examining the value of the multiplier in the case (considered by Keynes, GT, 122–123) where consumption-goods firms do not anticipate the increase in demand for their goods as a result of an increase in

employment in the capital-goods sector. OA considers two cases where temporary equilibrium is struck either by inventory adjustment at fixed prices or by market clearing prices. I will address both cases in turn.

<u>Inventory adjustment</u>

If it is assumed that cash received from the sale of inventory is not expended on consumption, OA notes that the reduction in the value of inventories equals the increase in fixed investment multiplied by the normal value of the marginal propensity to consume. So the equation reads:

$$\phi(N_A) - \phi(\tilde{N}) = \frac{\left[I_F^* - \bar{c}(I_F^* - \tilde{I}_F) - \tilde{I}_F \right]}{1 - \bar{c}} \tag{1}$$

This equation is consistent with the 'logical theory of the multiplier'. OA then argues that in the inventory adjustment case, if the cash received from the sale of inventory is directly or indirectly expended on consumption, the inventory reduction captured by the term

$$\bar{c}(I_F^* - \tilde{I}_F) \tag{2}$$

in the above equation will be increased, thus breaking the equality. He does not put a value on the revised inventory reduction but based on his diagram (Figure 4) it will be

$$\bar{c}(E^* - \tilde{E}) \tag{3}$$

However, the apparent refutation confuses receipts (which OA incorrectly calls 'money income') with income ('total income') (OA, 14–15). E^* represents the level of effective demand (total income, in OA's terms) when the multiplier has fully 'worked itself out'. OA is arguing that sales receipts permit consumption at the normal level consistent with E^*, so that the depletion of inventories equals the equilibrium increase in total income at the 'fully worked out' level, multiplied by the normal value of the marginal propensity to consume. This is quite possible, but it does not undermine the logical theory of the multiplier.

The refutation fails because Keynes's propensity to consume is defined as a relation between consumption and income, not between consumption and receipts. OA defines receipts (R) as equal to sales (S). Note that receipts (R) differ from income (Y) by the value of user cost (OA considers only inventory adjustment, I_L, since all inter-firm purchases are considered fixed investment, I_F, OA, 6n) but receipts (R) can also differ from sales (S) through borrowing. While Keynes finds it possible to postulate a function linking aggregate consumption to aggregate income (i.e. the value of aggregate output) without reference to changes in inventories or borrowing or the degree of integration of industry, no such function can exist between unique values of consumption and unique values of either sales or receipts. Certainly Keynes does not suggest the latter. Furthermore while a stable functional relation between consumption and income has some behavioural plausibility, this cannot be said for a link between consumption and sales or bank balances.

Thus what OA is capturing is the possibility of a temporary increase in the value of the marginal propensity to consume (out of *income*) which may be written as

$$c = \bar{c} \frac{(E^* - \tilde{E})}{(I_F^* - \tilde{I}_F)}$$

This is perfectly consistent with Keynes's statement (just preceding the quotation given by OA in a footnote to page 16) that 'so far as the balance is restored by a postponement of consumption there is a temporary reduction of the marginal propensity to consume, i.e. of the multiplier itself' (GT, 124)—albeit Keynes is in this quotation considering the opposite case. Indeed Keynes's presumption has greater behavioural plausibility since one would expect pressure on inventories to lead to price increases or shortages, either of which would tend to defer consumption. It is only OA's assumption of unimpeded quantity adjustment at fixed prices that leads to his particular permutation.

Whether or not receipts from sales of inventory are expended on consumption, the multiplier of the initial increase in fixed investment is unaffected and remains at unity since there is no employment response from the consumption industries in this case. Restating OA's equation (1):

$$\phi(N_A) - \phi(\tilde{N}) = \frac{\left[I_F^* - \overline{c}\left(I_F^* - \tilde{I}_F\right) - \tilde{I}_F\right]}{1 - \overline{c}} = \frac{\left(1 - \overline{c}\right)\left(I_F^* - \tilde{I}_F\right)}{1 - \overline{c}} \tag{4}$$

so that the multiplier can be written:

$$\frac{\phi(N_A) - \phi(\tilde{N})}{I_F^* - \tilde{I}_F} = \frac{1 - \overline{c}}{1 - \overline{c}} = \frac{1 - \overline{c}\left(E^* - \tilde{E}\right)/\left(I_F^* - \tilde{I}_F\right)}{1 - \overline{c}\left(E^* - \tilde{E}\right)/\left(I_F^* - \tilde{I}_F\right)} = 1 \tag{5}$$

Price adjustment

In section 6 of his paper, OA considers adjustment to equilibrium through price adjustment during market clearing at the end of each elementary period. There is therefore no change in inventories so by definition $I_L = 0$ and $Y = R$ (OA, 20n).

The text is a little confusing here since it refers to states of expectation $e_A{}'$ and $e_A{}''$ which appear to correspond to e_B and e_C in Figure 5. Nevertheless, since in this case there is no difference between R and Y, OA finds in support of Keynes that the logical multiplier holds both in temporary and stationary equilibrium.

In summary, OA's purported correction of Keynes and refutation of the logical theory of the multiplier depends on a redefinition of the consumption function into something that cannot be uniquely defined and lacks behavioural plausibility.

REFERENCES

Allain, O. 2009. Effective demand and short-term adjustments in the General Theory, *Review of Political Economy*, vol. 21, 1–22

Hartwig, J. 2007. Keynes vs. the Post Keynesians on the principle of effective demand, *European Journal of the History of Economic Thought*, vol. 14, 725–39