Keynes’s Aggregate Supply Function: everything you always wanted to know about Z

Jochen Hartwig
Keynes’s Aggregate Supply Function: everything you always wanted to know about Z

Jochen Hartwig  
KOF Swiss Economic Institute at ETH Zurich  
Weinbergstr. 35, 8092 Zurich  
Switzerland  
Telephone:  +41-44-632-7331  
Fax:  +41-44-632-1218  
Email:  hartwig@kof.ethz.ch

JEL classification:  B31, C43  
Keywords:  Keynes, aggregate supply function, output heterogeneity

Abstract  
The paper suggests a consistent interpretation for the much debated Z-footnote on pp. 55-56 of the General Theory and discards claims recently made in the literature concerning the importance of output heterogeneity for Keynes’s macroeconomic approach.

1. Introduction  
Something’s wrong with footnote 2 on pp. 55-56 of Keynes’s General Theory, so it seems. In that footnote, Keynes makes two seemingly contradictory statements about the slope of the aggregate supply function Z. First he says that the slope of Z in wage units (Z_w) is equal to 1. In the last sentence of the footnote however, he states that the slope of the aggregate supply function is equal to the reciprocal of the money wage.

In a recent debate on the footnote both Hayes (2007, 2008) and Hartwig and Brady (2008) were able to make sense of the first part of the footnote by showing (albeit in quite different ways) that the slope of Z_w equals one. Also, both Hayes and Hartwig and Brady came to the conclusion that the last sentence of the footnote must be wrong. Gerhard Michael Ambrosi has now joined the debate, advertising his interpretation “as a candidate for settling the long
dispute over ‘what Keynes meant’” in the footnote (Ambrosi, 2011, p. 633). Ambrosi in principle agrees that the last sentence is wrong but claims that it could be corrected by inserting the word ‘share’ at the very end of the footnote. In other words, Ambrosi thinks that Keynes ‘really meant’ that the slope of the aggregate supply function is equal to the reciprocal of the money wage share.

In this paper I will argue that although Ambrosi’s contribution indeed pushes the debate forward into the right direction, there is still room for improvement and clarification. Ambrosi’s claim that the last sentence of the footnote has something to do with the inverse of the wage share is indeed correct, as is his insistence that the first part of the footnote and the last sentence deal with different curves. In other respects, however, there are a number of misrepresentations and even false claims in Ambrosi’s article which need to be rectified. One such misrepresentation concerns his emphasis on the heterogeneity of output.

2. **Heterogeneous output: does it matter for macroeconomics?**

Ambrosi joins Hayes (2007, 2008) in emphasizing the heterogeneity of output and the alleged inadmissibility of dealing with concepts such as aggregate output or the economy-wide production function.¹ Hartwig and Brady (2008), on their part, while acknowledging that Keynes expresses his unease with concepts like aggregate levels of prices and output in chapter 4 of the *General Theory*, plead for common sense. They argue that it would be unwise to drop a useful construct like real GDP from macroeconomic analysis. Ambrosi clearly misrepresents their argument when he insinuates that they are assuming a central planning bureau instead of acknowledging that the economy consists of a multitude of profit-maximising firms. In earlier papers on Keynes’s principle of effective demand – see Hartwig (2007) for just one example – I repeatedly addressed the question how the macroeconomic model of supply and demand from chapter 3 of the *General Theory* can be construed taking profit maximisation of individual firms as starting point.² This is a non-trivial problem,

---

¹ Apparently, however, Ambrosi and Hayes draw different conclusions for the aggregation of firms’ heterogeneous output. Ambrosi (2011, p. 628) arrives at the macroeconomic Z function by summing firms’ individual Z functions. Hayes, on his part, writes: “It is natural to think of the aggregate supply function being built, from the bottom up as it were, as the sum of the aggregate supply functions of individual firms: based upon the ‘microfoundation’ of industry supply and demand. … Keynes’s method runs the other way, from top down, from the aggregate supply function of industry as a whole, to those individual industries: truly ‘macrofoundation of microeconomics’” (Hayes, 2007, p. 744).

² Ambrosi, in his footnote 8, displays a curious understanding of the term ‘macroeconomic’ when he claims that only one-sectoral models of the overall economy are macroeconomic models. This is of course not the case. Also
especially for the aggregate demand function \((D)\), in my opinion. For the aggregate supply function, \(Z\), however, I tend to share Ambrosi’s (rather than Hayes’s) view that the individual firms’ \(Z\) functions can be summed to arrive at the macroeconomic \(Z\) function. Of course one must measure the individual firms’ outputs in terms of money in order to be able to aggregate them. But once we have summed the firms’ sales and deducted their purchases of intermediate inputs to arrive at nominal GDP so to speak, we can deflate this magnitude appropriately in order to separate the price and the quantity component in GDP or, more precisely, the change in these two components from one period to the next. Then it is also possible to construct an index representing the overall price level – which has the value of 100 in an arbitrarily chosen base period – and an index representing the overall level of real output – which has the same value as nominal GDP in the base period. This is standard practice in statistical offices around the world. Nowhere along the way do we need to assume a central planning agency or that every firm produces the same product. No-one denies that output is heterogeneous in the real world. The point is that we can construct a homogenous measure for the aggregate of heterogeneous real output with which we can do macroeconomic analysis.

Froyen (1976), in his study on the aggregative structure of the *General Theory*, draws attention to the fact that Keynes’s main concern was explicitly with “the Theory of Output and Employment as a whole” (Keynes, 1936, p. 293). Froyen (1976, p. 381) concludes that considerations concerning “the contrast … between the one-good model and the \(n\)-good economy the model represented … were of secondary importance” for Keynes because he ignored them in the formal model of the *General Theory*, which consists of only four aggregates: commodities, money, labour services and nonmoney assets, and three relative prices: the level of commodity prices, the interest rate, and the money wage. As a matter of fact, Keynes uses the concepts of aggregate real output \((O)\) and the aggregate price level \((P)\) several times in the *General Theory*. Normally, he adds ‘warning notices’ such as: “if it is practicable to measure the quantity, \(O\), and the price, \(P\), of current output” (Keynes, 1936, p. 209) or “(i)f we are dealing with industry as a whole and are prepared to assume that we have a unit in which output as a whole can be measured” (Keynes, 1936, p. 285). In the remainder of this paper, I will assume that this is indeed practicable. Following Chick (1983, p. 66), I a two-sectoral model (like, for instance, Pigou’s model which Keynes discusses in the appendix to chapter 19 of the *General Theory*) is a macroeconomic model and the corresponding production functions of the wage goods and non-wage goods sectors are macroeconomic production functions as opposed to microeconomic production functions of single firms.

\footnote{Let us not forget that at the time Keynes wrote the *General Theory*, National Accounting did not yet exist. As a matter of fact, Keynes’s theory was an important inspiration for economists like Nobel laureates James Meade...}
will therefore assume that “Z(N) is derived from the aggregate production function” \( O = \psi(N) \), keeping in mind that this relies on a number of critical assumptions, the realism of which may be disputed.4

3. Explaining the Z-footnote
Hartwig and Brady (2008) explain the first part of the Z-footnote – the proposition that the slope of the \( Z_w \)-curve equals 1 – as follows: \( \frac{dZ_w}{dN} = 1 \) can be transformed into \( \frac{dZ}{dN} = w \).

Integrating with respect to \( N \) yields:

\[
\int \frac{dZ}{dN} dN = \int w \, dN \iff Z = w \cdot N + C
\]

Since \( Z \) equals the sum of wages and expected profits (see Hartwig and Brady, 2008), the first part of the footnote would make sense if the constant of integration \( C \) denoted expected profits. However, normally profits are not constant along \( Z \); they are rising with \( N \).5 What is constant along a linear \( Z \) curve is the profit share – and therefore, of course, also the wage share – because wage income is also a linear function of \( N \) in the \( Z/N \)-diagramm (see Chick 1983, p. 96, for instance). This suggests that Ambrosi’s conjecture that the footnote has something to do with income shares is indeed correct.

This can be shown much more straightforwardly than Ambrosi has done, however. The major flaw in Ambrosi’s exposition is that he denies that Keynes, in the first part of the footnote, actually states that the slope of \( Z \) in wage units (\( Z_w \)) is equal to 1. But Keynes defines “\( Z_w = \phi(N) \)” and concludes: “\( \phi'(N) = 1 \)”. According to the generally accepted mathematical notation, this means that the derivative of \( Z_w \) with respect to \( N \), i.e. the slope of the \( Z_w \)-function in a \( Z_w/N \)-diagramm, equals 1. Ambrosi disputes the indisputable when he writes: “Let us stress once more that the unitary value of the slope is not a statement about the

and Richard Stone who began developing the System of National Accounts during World War II (see Abrahamsen and Hartwig, 2007). One might conjecture that Keynes’s warnings against measures of aggregate output in the General Theory would have been less pronounced if he had already known the achievements of modern National Accounts.

Chick (1983, p. 65f.) mentions the following three assumptions: “(i) atomistic firms; (ii) that labour is the only variable factor of production …; and (iii) that the composition of output and demand does not change with the overall volume of output”.

There is a special case in which profits are indeed constant along \( Z \), and I will argue below that the first part of the footnote targets exactly this special case.
shape of the $Z_w$-function itself” (Ambrosi, 2011, p. 630). His compelling argument that the last sentence of the footnote states something about the wage share must therefore be derived differently than he has done. It must be derived in a way that respects the wording of the first part of the footnote.

Another of Ambrosi’s insights may be instrumental in doing so, namely the notion that the first part of the footnote and the last sentence deal with different curves. He writes that the footnote “refers to characteristics of every individual point which lies on this [aggregate supply] curve but which might at the same time be seen as a point of a different curve – namely of a 45°-curve – which cuts the $Z_w$-curve in every single point of the $Z_{wr}$-curve” (Ambrosi, 2011, p. 621). To understand this somewhat obscure sentence, it is probably necessary to have a look at Ambrosi’s Figure 1, which has a quadrant with a production function – of course only for the individual firm because Ambrosi denies the admissibility of the aggregate production function – a $Z_w$-curve (also with subscript ‘$r$’ indicating the microeconomic level of analysis) and a 45°-line which mirrors points on the production function onto the $Z_{wr}$-curve. Now Ambrosi believes that this 45°-line is the function meant by Keynes in the first part of the footnote, and he speaks of ‘Keynes’ 45°-line’ (although Keynes in the footnote never mentions a 45°-line; he just states that the slope the $Z_w$-function – without ‘$r$‘ subscript, one might add – equals 1).

Although Ambrosi’s exposition is hence inadequate, his insight that the two parts of the $Z$-footnote deal with different curves is valuable. It has to be pointed out, however, that this insight has already been published in early 2009 by Claudia Heller (Heller, 2009), who, on her part, builds on Dos Santos Ferreira and Michel (1991). Ambrosi does not mention these contributions. Since Heller’s account of the issue is much clearer and more in line with the wording of the footnote than Ambrosi’s, it will therefore be briefly restated.6

Heller points out that Keynes in the footnote distinguishes between an ‘aggregate supply function’ $Z_w = \phi(N)$ and an ‘aggregate supply curve’. This distinction is so subtle that most interpreters – including myself – have overlooked it so far, believing that the two expressions signify the same thing. Heller, however, draws attention to p. 44 of the General Theory, where Keynes already uses the expression ‘supply curve’, which – ignoring ‘$r$’ subscripts and ‘user cost’ – he defines as:

$$p = \frac{Z}{O} = \frac{\phi(N)}{\psi(N)}$$  \hspace{1cm} (2)

---

6 Dos Santos Ferreira and Michel’s (1991) focus is not the footnote so I will go along with Heller. With regards to content, the former reference would be equally adequate.
So the ‘aggregate supply curve’ is not the same thing as the ‘aggregate supply function’ \((Z)\). Rearranging (2) yields:

\[
p \cdot O = Z \tag{3}
\]

This means that, in terms of chapter 3 of the *General Theory*, the ‘aggregate supply curve’ consists of conceivable ‘points of effective demand’, i.e. of points of intersection between the ‘aggregate supply function’ \(Z\) and the ‘aggregate demand function’ \(D = p \cdot O\), for different conceivable demand price levels \(p\).\(^7\) Dos Santos Ferreira and Michel (1991) and Heller (2009) denote this \((Z = D)\)-curve – in other words, the ‘aggregate supply curve’ – as \(Z^*\). Both Dos Santos Ferreira and Michel (1991) and Heller (2009) acknowledge a resemblance of this curve to the ‘employment function’ of chapter 20 of the *General Theory*, which Keynes also defines as a collection of points of effective demand.

Now, as Keynes makes clear in chapter 3 of the *General Theory*, each point of effective demand entails profit maximisation. This implies, as he reiterates in the \(Z\)-footnote, that the marginal product must be equal to the marginal factor cost along \(Z^*\). This means that the price component implicit in \(Z^*\) – which we can call \(P^*\) – must be equal to the wage rate divided by the marginal product of labour. We therefore know that the \(Z^*\) and the \(Z_w^*\) curves, respectively, are given by:

\[
Z^* = \frac{W}{\psi '(N)} \cdot \psi (N) \iff Z_w^* = \frac{\psi (N)}{\psi '(N)}
\tag{4}
\]

What we really want to know is the slope of \(Z_w^*\). Remembering that \(Z^*\) equals \(P^* \cdot O\), we know that \(Z_w^*\) equals \((P^*/W) \cdot O\) or \(P_w^* \cdot O\). Then we can write (see Arthmar and Brady, 2009, Appendix 3):

\[
\frac{d(P_w^* \cdot O)}{dN} = \frac{d(P^* \cdot O)}{dN} \cdot \frac{1}{dO/dN} \cdot \psi'(N) = \frac{d(P^* \cdot O)}{dO} \cdot \psi'(N)
\tag{5}
\]

Knowing that \(\psi'(N) = W / P^* = 1 / P_w^*\) because points on \(Z^*\) are profit-maximizing, we can rewrite (5) as:

\[
\frac{d(P_w^* \cdot O)}{dO} \cdot \psi'(N) = \frac{d(P^* \cdot O)}{dO} \cdot \frac{1}{P_w^*}
\tag{6}
\]

\(P_w^* \cdot O\), the effective demand in wage-units, Keynes calls ‘\(D_w\)’ in chapter 20 of the *General Theory*. So we can rewrite (6) as:

\[
\frac{d(P_w^* \cdot O)}{dO} \cdot \frac{1}{P_w^*} = \frac{dD_w \cdot O}{dO} \tag{7}
\]

\(^7\) For the distinction between the supply price level and the demand price level, see Hartwig (2007), for instance.
On the right-hand side of equation (7) we have the inverse of what Keynes (1936, p. 282f.) defines as ‘the elasticity of output or production’. This is therefore the most general expression for the slope of the ‘aggregate supply curve’: it equals the inverse of the output elasticity. Under conventional (neo)classical supply-side assumptions, which Keynes accepted in the General Theory – I agree with Ambrosi on this point – the inverse of the output elasticity equals the inverse of the wage share. So Ambrosi’s conjecture on the last sentence of the footnote is right, but the result can be derived in a straightforward way, as equations (2) to (7) show.

To sum up, there are two errors in the last sentence of the footnote: (i) instead of ‘aggregate supply function’ it should read ‘aggregate supply curve’, and (ii), with respect to the slope of the curve, it should read ‘reciprocal of the money wage share’ – or ‘reciprocal of the output elasticity’ – instead of ‘reciprocal of the money wage’.

Finally, we can – at least partially – reconcile the last sentence of the footnote with its first part. “(I)f the elasticity of output is unity”, writes Keynes (1936, p. 283), “no part of the increased effective demand is expected to accrue as profit”. From solving equation (1) above backwards, however, we know that in this special case of zero marginal profits, the slope of the ‘aggregate supply function’ in wage-units ($Z_w$) equals 1. So the statement of the first part of the footnote is correct, but only for the special case of an output elasticity – and hence a wage share – of 1.

3. Conclusion
Ambrosi (2011) has joined the recent debate over Keynes’s Z-footnote on pp. 55-56 of the General Theory. He adds two important insights: (i) that the word ‘share’ has to be inserted at the very end of the footnote in order to make sense of Keynes’s statement about the slope of the aggregate supply curve, and (ii) that the first part of the footnote and its last sentence are concerned with different functions or curves. However, the originality of (ii) is questionable; and Ambrosi’s treatment of the first part of the footnote is in contradiction with Keynes’s explicit statements about the aggregate supply function in wage-units. (i), on its part, can be derived more rigorously than Ambrosi did. To point this out was my first aim with this paper. My second aim was to fight a tendency that otherwise threatens to take root, namely the tendency to use the notion of heterogeneous output in order to deprive macroeconomics of some of the most important and useful concepts it has developed so far. Real GDP is an example for such a concept that scholars like Ambrosi deem inadmissible.
Bibliography


Heller, C. 2009. Keynes’s slip of the pen: aggregate supply curve vs. employment function, MPRA Paper No. 12837