Teaching marginal analysis: On the importance of emphasising the second-order condition

Akihito Asano

Abstract

Students employ marginal analysis in investigating various problems regarding resource allocation. The majority of first-year students, however, seem to struggle in applying it to a firm's profit maximisation problem. Their confusion appears to stem from textbooks for introductory microeconomics courses oversimplifying some subtle points regarding the profit maximisation problem. This paper (1) points out how oversimplification may impede students' understanding, contrary to the instructors' pedagogical intent, and (2) suggests a complementary method to enhance their understanding of marginal analysis in the course of teaching a firm's profit maximisation problem.

Introduction

Marginal analysis is the heart of economics. It is one of the most important concepts students learn throughout their studies. In teaching undergraduate students this important concept, we are conscious that whilst we know so much about it the students know only little. Consequently we tend to fudge some subtle issues for pedagogical reasons especially in the first year courses. Economists who write introductory microeconomics textbooks are no exception. These textbooks tend to hide subtle issues that do not seem to be essential to the focus of the problems.[1] Presumably it is done in order to to avoid students being confused by matters that are nonessential. However, hiding the subtlety sometimes may have a negative pedagogical effect. Students tend to think in ways instructors never expect they would, or they tend to get stuck with points instructors never expect

they would. If students are stuck with marginal analysis in an introductory microeconomics course, it is a big problem for the instructors who teach those students in later-year units. It is a problem also in terms of enrolment if those students give up studying economics.

In order to illustrate the above problem in teaching marginal analysis, we focus on the firm's profit maximisation problem. In a typical introductory microeconomics textbook, this topic starts with an explanation of a production function with one variable input, say labour, and the associated cost function. In a typical production function in these textbooks, marginal product increases when quantity is low, but as quantity rises diminishing marginal product kicks in. This obviously implies that the marginal cost (*MC*), the average cost (*AC*), and the average variable cost (*AV C*) curves are all U-shaped. Then the textbooks analyse a firm's behaviour in a competitive market where marginal revenue (*MR*) is constant at the market price of the commodity. Marginal analysis is then employed to determine the optimal level of output. When *MR* is greater than *MC*, the firm should produce an extra unit of the commodity. It should stop producing when *MR* and *MC* are equal. Therefore, given the market price of the commodity *p*, the firm will choose the optimal quantity *q*, such that $p = MC(q)$.[2] In a competitive market, this is how the supply curve for a firm is derived.

The majority of the students, however, seem to have trouble in fully understanding the above simple flow of argument. In the following section, we endeavour to explain typically how textbooks for introductory microeconomics treats this material and why that could be a pedagogical problem. Section 3 proposes a method to effectively teach marginal analysis in the context of a firm's profit maximisation problem. The method is free from the use of calculus. Section 4 summarises our argument using calculus, which helps more mathematically inclined first-year students consolidate their understanding on this matter. Section 5 concludes the paper.

Problems

We illustrate relevant problems using a typical numerical example that students might face in an introductory microeconomics course.

Question: The table attached provides cost information of a firm supplying a commodity in a competitive industry (see Table 1). The technology of the firm is such that $AC(q)$, $AV\,C(q)$, and $MC(q)$ curves are all U-shaped. *What is the optimal level of output if the price of this commodity is $140? Obtain the profits.*

Table 1 Data for various costs

q	*VC(q)*	*TC(q)*	*MC(q)*	*AV C(q)*	*AC(q)*
0	0	1000	NA	NA	NA
1	181	1181	181	181	1181
2	328	1328	147	164	664
3	447	1447	119	149	482
4	544	1544	97	136	386
5	625	1625	81	125	325
6	696	1696	71	116	283
7	763	1763	67	109	252
8	832	1832	69	104	229
9	909	1909	77	101	212
10	1000	2000	91	100	200
11	1111	2111	111	101	192
12	1248	2248	137	104	187
13	1417	2417	169	109	186
14	1624	2624	207	116	187
15	1875	2875	251	125	192

The answer is clearly $q = 12$ and the corresponding profit is –$568 (the loss of $568 which is less than the fixed cost of $1000).[3] However, many students surprisingly tend to choose $q = 0$. Our concern becomes even greater when we find that the major reason for them choosing $q = 0$ is not because they obtain the negative profit of –$568 when q =12. These students do not even arrive at the conclusion $q = 12$. The most typical error is as follows: *To produce the first unit, it costs $181 which is less than the price, therefore the firm should not produce at all. The corresponding profit is –$1000.*

This is a big concern. Many students clearly do not understand (or are confused by) marginal analysis. The objective of this section is to detect problems that lead to this misunderstanding. There seems to be one major problem, which is explained next. We then raise a related problem and explain that these two problems combine together to create confusion amongst students and impede their understanding.

Fudging

As mentioned in the previous section, production theory in introductory microeconomics typically starts with concepts of marginal product of an input. A production function, in which marginal product increases and diminishes, is introduced followed by the associated cost function. This implies that *MC*, *AC* and *AV C* curves are all U-shaped. Then the relative positions of these three curves are explained, i.e. when *AV C* is above *MC*, *AV C* is decreasing, etc.

With U-shaped curves, we face an awkward problem. When we introduce marginal analysis and arrive at the conclusion that, given the market price of the commodity *p*, the firm will choose the optimal quantity *q* such that $p = MC(q)$, there are two possible levels of quantity that satisfy this condition. The majority of the introductory microeconomics textbooks seem to fudge this problem in one of the following ways.

Fudging 1 *The downward sloping segment of the MC curve is simply omitted (Figure 1).*

Some textbooks depict the *MC* curve as in Figure 1. Despite the fact that the emphasis is given to both increasing and diminishing marginal products, the *MC* curve is only increasing. By drawing the *MC* curve like this, we can obviously avoid the problem of having two solutions to $p = MC(q)$. That is, by starting with some arbitrary taken quantity q_0 in Figure 1 where $p > MC(q_0)$, it is quite straightforward to arrive at q where $p = MC(q^*)$.[4]

Later in this section we shall discuss why this fudging is problematic, but before doing that let us review other ways in which textbooks fudge this matter.

Figure 1 Decreasing segment of MC(q) is ignored

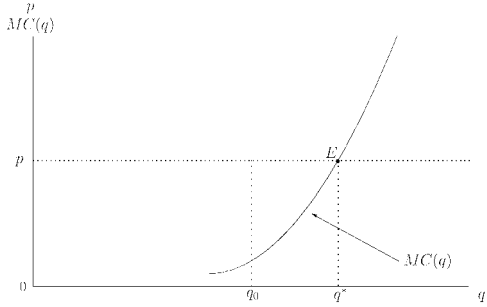
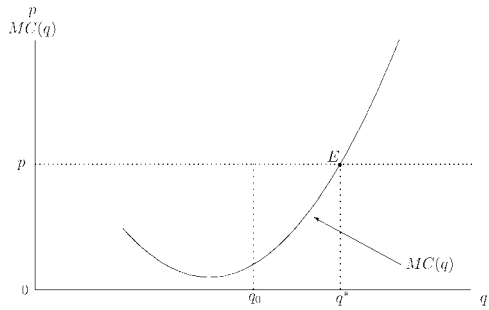


Figure 2 Still fudging



Fudging 2 *The downward sloping segment of the MC curve is drawn, but only partly so that p = MC(q) occurs only at one q (Figure 2).*

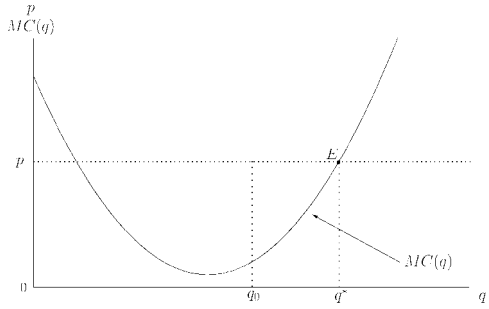
Is this a good compromise? Whilst the *MC* curve is U-shaped, there is only one point on the diagram such that $p = MC(q)$. But it is only so because a part of the *MC* curve is still ignored. So there is not much difference between these two ways of fudging.

Some textbooks do draw that segment of the *MC* curve so that there are two intersections, but still fudge this problem.

Fudging 3 *The MC curve is U-shaped and p = MC(q) occurs at two levels of q, but one of those is completely ignored and is not even labelled (Figure 3).*

In a sense, this is better than the previous two ways of fudging as the *MC* curve is at least consistent with the production technology that is emphasised. However, even though it has now revealed there are two intersections, without commenting at all about the unlabelled intersection, textbooks focus on the other intersection and move on to derive the short-run supply curve, which we know is upward sloping.

Figure 3 No attention to one of the intersections



Memorising

There seem to be a related problem regarding students' attitude towards studying, which teachers need to take into account when they deliver lectures. The problem is that most of the students are weak in logic and tend to accept statements as gospel.

Let us put this in the context of the profit maximisation problem. Some textbooks emphasise the fact that a competitive firm's *MC* is increasing at the profit maximising level of output (i.e. suggesting to students that they can ignore the downward sloping segment of *MC*), but it seems to have little effect on enhancing students' understanding. In addition, it can be dangerous to provide students with this sort of statement without emphasising reservations.[5]

Fudging 4 *A firm produces an output where MR equals MC and, at this output, MC is increasing.*[6]

This statement, which sometimes seems to be referred to as the *golden rule of profit maximisation*, is true, but not always. It is true when the firm is a price taker (i.e. competitive markets). A firm can increase its profit whenever *MR* is greater than *MC* by increasing its production. Therefore, *in a competitive market* where *MR* is constant at the market price of the commodity, the only way for the profit-maximising firm to stop increasing its production is to have an upward sloping *MC* curve. Instructors are aware of this, but a typical first-year student tends to take this statement as gospel and believes that it holds for any type of market. That is, they tend to just memorise the last part of the statement: *MC* is increasing (at the profit maximising point).

Discussion: What is the problem?

Readers might want to ask two questions at this point. One might be, 'In introductory microeconomics, doesn't it suffice to restrict the attention to a *MC* curve that is always upward sloping even though it contradicts the typical production technology that has been introduced?' The other might be, 'If students take the 'golden rule' as gospel then why do they choose a quantity where *MC* is decreasing ($q = 0$) in the numerical example we saw previously? 'The first question was in fact raised in the beginning of this section. The reason we have postponed answering it is that it can be better answered if the second question is taken into account at the same time.

The first question *per se* appears to be debatable. Do we worry about the fact that a *MC* curve is inconsistent with the typical production technology, when our focus is explaining marginal analysis in the context of firm's profit maximisation problem? If not, does this inconsistency confuse students? This perhaps is a minor problem.

Students will accept a *MC* curve that is always upward sloping without difficulty if an instructor acknowledges it.[7] If this is not a problem, then why bother worrying about the downward sloping segment of the *MC* curve when we know that no solution can be found there, and hence is inessential, in a perfectly competitive market?

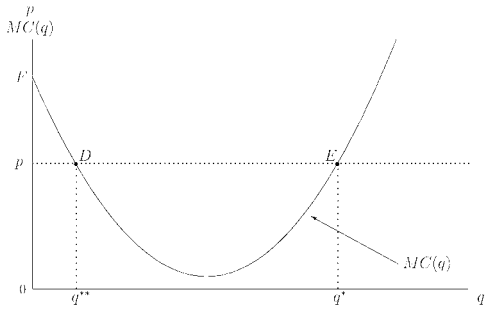
It is true that the downward sloping segment of the *MC* curve is nonessential in the sense it is not going to be a solution in a perfectly competitive market. However, we would like to argue here that emphasising that segment is essential in terms of pedagogy on the contrary to what one might think.

We choose not to teach various things in introductory microeconomics for pedagogical reasons. Monopoly with a kinked demand curve is one of these examples.[8] We do not usually teach it in introductory economics because our major objectives at this level are to teach what monopoly is and to demonstrate the inefficiency that it causes – marginal analysis, therefore, is a part of it – in the market.[9] Some technicality regarding the kinked demand is left for intermediate microeconomics or industrial organisation. In this case, hiding the kinked demand does no harm to students' understanding of efficiency, so it is pedagogically nonessential to emphasise that peculiar demand. On the contrary, however, in the firm's profit maximisation problem that we have focussed, emphasising the downward sloping segment of the *MC* curve is essential. This is because marginal analysis is one of the main things students are supposed to learn from going through that problem.

Considering the answer to the second question makes the point clear. Recall that the question was, 'If students take the 'golden rule' as gospel then why do they choose a quantity where *MC* is decreasing ($q = 0$) in the numerical example we saw previously? 'The reason is because the fact that *MC* is decreasing does not matter to them. They have decided that the firm does not produce on the basis of $MR < MC$ for the first unit, and that is the end of their reasoning. They do not check the slope of *MC* because they do not *appreciate* the importance of it. Students are not exposed to the reasoning as to why it cannot be profit maximising when *MC* is downward sloping in perfect competition, and a typical student is not so logical to be able to figure out that profit maximising cannot occur at that segment of the *MC* curve.

As a consequence, a typical student results in (1) memorising that *MC* is increasing when a competitive firm is profit maximising but; (2) still looking at the downward sloping part of the *MC* curve when they face a numerical problem. The problem is a combination of (1) omitting explanation as to why one of the intersections is not profit maximising and; (2) students trying to resort to memorising concepts rather than to understand them logically.

Figure 4 Two points satisfy *MR* = *MC*



We therefore believe that emphasising the downward sloping segment of the *MC* curve has a huge pedagogical effect, both in terms of instilling students how marginal analysis works and developing their logical thinking.

One way it could be taught

Taking our discussion in the previous section into account, we shall suggest a way the profit maximisation problem could be taught in the following.

Step 1 *Draw a U-shaped MC curve and acknowledge that MR(q)= MC(q) occurs at two levels of q in perfect competition. Show that one of the two points might be profit maximising.*

At the same time, teachers need to explain first that *MR(q)= p* for all *q*, hence the *MR* curve is constant at *p*. The important thing from now on is to keep using *MR* instead of *p*. Students just need to be persuaded that *MR* is constant in perfect competition.

Step 2 *Show that the other point cannot be profit maximising.*

This is perhaps the most important step, which the textbooks tend to omit. By the use of *reductio ad absurdum*, it can be taught why Point *D* cannot be profit maximising, and indeed this re-emphasises the power of the marginal analysis. Suppose that Point *D* were indeed profit maximising. Then by marginally increasing its production, the firm would be able to increase its profit as *MR(q) >MC(q)*. This contradicts to the assumption that Point *D* is profit maximising, and hence it cannot be profit maximising. It is very important to emphasise here that when the firm is maximising its profit, at that point, the *MR* curve cuts the *MC* curve from *above*. This condition holds at Point *E* but not at Point *D*.

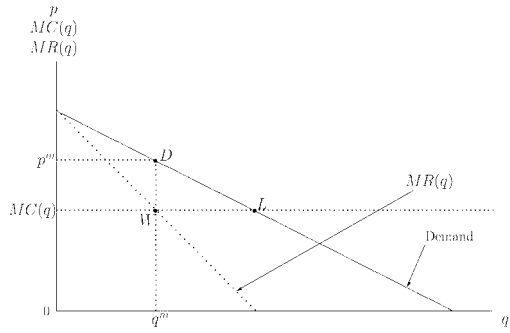
In any event, this leaves a firm with a decision of producing at Point E or not producing at all (Point F). Obviously, choosing Point F results in a negative profit equivalent to the fixed cost (FC) hence when Point E is above the AVC curve, the firm will choose to produce. Otherwise, the firm decides not to produce.

Step 3 *Keep using MR(q) instead of the market price. Teach monopoly before deriving an upward sloping supply curve in perfect competition.*

In the previous section, we have acknowledged the danger of just emphasising the upward sloping segment of the *MC* curve. Students tend to take it as gospel that the *MC* curve is upward sloping for any profit maximising point, which is true if perfect competition is assumed but it is not necessarily true in general. What we want to stress is that what generally matters is the difference between slopes of *MR* and *MC*, but not the slope of *MC* itself. We had better instill the notion of marginal analysis into the students using the two different markets, perfect competition and monopoly, back to back. Typically in the first year, we teach the concept of natural monopoly, yet the *MC* curve tends to be drawn upward sloping when the optimal output for monopoly is derived. This merely puts off students' (possible) misunderstanding on this matter to later years.

Instead it is suggested that we use natural monopoly, which is usually explained in the beginning of a monopoly section in introductory microeconomics textbooks. Assuming that monopoly has a total cost function $C(q)= F + cq$ where F and c are the *FC* and the *MC*, respectively, we can explain all the important points including the deadweight loss. With this cost function and linear demand, profit maximisation for monopoly occurs at q^m in Figure 5 where $MR(q^m)= MC(q^m)$.[10] Students just need to apply the same principle of the marginal analysis once they figure out the *MR* and *MC* curves.

Figure 5 Monopoly: The same principle applies



There are two important points to emphasise in this example. First is that at Point *W* the *MC* curve is non-increasing even though a monopoly is profit maximising. This example shows that the *MC* curve is not necessarily upward sloping at a profit maximising point. The second point to emphasise is that the *MR* curve is not constant, more specifically, it is downward sloping. This is what introductory microeconomics textbooks explain intuitively (in relation to elasticity), but in this paper, the focus is to technically explain that the *MR* curve is not constant and $p>MC(q^m)$ at the profit maximising quantity (which results in the deadweight loss).

Step 4 *Come back to perfect competition and derive an upward sloping supply curve. Emphasise that when the market is competitive, i.e. when a firm is a price taker, MR(q) is constant at the market price p.*

We have emphasised that we always have $MR(q)= MC(q)$ for profit maximisation and that the *MR* curve cuts the *MC* curve from above at the optimum. Now we re-emphasise to students that perfect competition is the special case where $MR(q)$ is always p. We go through this because we want to demonstrate to students that we can derive an upward sloping supply curve for a price-taking firm. Realising that there is rigorous foundation to the law of supply, students will, without doubt, become more enthusiastic in studying economics.

Consolidating the understanding using calculus

Technically speaking, the focus of our discussion is, 'Do we teach the second-order condition of the profit-maximisation problem or not?' As we have seen, in introductory microeconomics textbooks, the first-order condition $MR(q) = MC(q)$ is emphasised but little attention is given to the second-order condition $MR'(q) <MC'(q)$. We have emphasised that we should give more emphasis on the latter.

Up to this point, we avoided the use of calculus bearing in mind that the material is for an introductory microeconomics course. For classes that are more mathematically inclined, however, the use of calculus might help consolidate students' understanding. What follows entails important points emphasised in the previous section. The use of calculus also highlights the fact that our idea is general and hence is applicable to different market structures.

Let us denote revenue and total cost by $R(q)$ and $C(q)$, respectively. Then the first-order condition for profit maximisation is:

$$R'(q)= C0(q), \tag{1}$$

which merely says that the optimum requires *MR* and *MC* to be equal. Whilst this condition holds in any structure of markets, in a competitive market, since $p = R0(q)$ it reduces to:

$$p = C'(q), \tag{2}$$

which just says a competitive firm produces a quantity where the price equals the marginal cost.

Now, the second-order condition for a local maximum is:

$$R''(q) < C''(q), \tag{3}$$

or alternatively,

$$MR0(q) < MC0(q). \tag{4}$$

Again, this is a general condition that applies to any market structures. In a competitive market, since $R'(q) = 0$ it boils down to:

$$MC'(q) > 0, \tag{5}$$

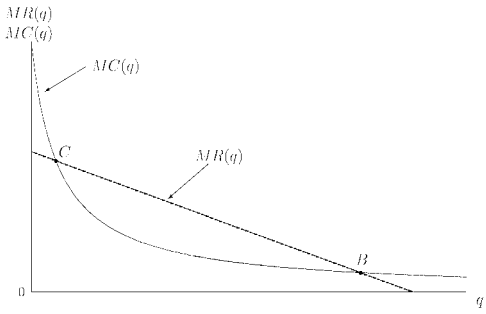
which says that the *MC* curve is upward sloping at the local maximum.

One of the points of the teaching method introduced in the previous section was to explain monopoly using Equation (4) first, and then come back to perfect competition, which uses Equation (5). Equation (4) clearly shows that what is important is the difference in slopes of *MR* and *MC* curves, not whether the MC curve is upward sloping or not.

The following example illustrates the point. Look at Figure 6. In this diagram, we have a monopoly who faces a linear demand curve. Hence its *MR* curve is linear as well. Suppose it has a decreasing *MC* curve as depicted. Equation (4) suggests that Point *B* satisfies the second-order condition but Point *C* does not. Indeed, if Point *C* were optimal, the monopoly should not be able to increase its profit by infinitesimally increasing the level of production. However, since the *MC* curve cuts the *MR* curve from above at Point *C*, the monopoly can increase its profit by increasing production. Hence Point *C* cannot be profit maximising for the monopoly.

Note that even at Point *B*, the *MC* curve is decreasing, i.e. $MC'(q) = C''(q) < 0$. In fact,this is true for all $q > 0$ in this example. Still, Point *B* is profit maximising as the *MR* curve cuts the *MC* curve from above and the monopoly cannot increase its profit by infinitesimally changing its production from there. This example helps emphasise the fact that what really matters in the profit maximisation problem is

Figure 6 Point *B* satisfying the second-order condition



the difference in slopes of *MR* and *MC*. The slope of the *MC* curve need not be positive at the profit maximising point. Students who have taken the 'golden rule' as gospel are likely to be confused when they see this because they would have just memorised that the *MC* has to be upward sloping when profit is maximised.

Concluding remarks

This paper is motivated by anecdotal evidence – a numerical example examining a competitive firm's profit maximisation problem – that suggests that the majority of first-year students do not fully grasp the concept of marginal analysis. Economists all know that understanding marginal analysis is crucial in studying economics. Understanding this material may be pivotal for students in determining whether or not to continue studying economics. If they do not believe in marginal analysis, there is no way they will continue to pursue their careers in economics, so we ought to investigate how we can teach this material effectively.

We have emphasised the fact that introductory microeconomics textbooks are omitting a subtle issue – the second-order condition – in a firm's profit maximisation problem. It may seem fudging the second-order condition of profit maximisation is justifiable especially when an introductory microeconomics course is concerned. When we know that the downward sloping segment of the *MC* curve is not essential to the problem in the sense that any point of that segment cannot be optimal, then why bother worrying about it? This must be the popular view considering how the textbooks treat this matter. Our thesis in this paper, however, is that we should not fudge teaching the second-order condition as it is pedagogically essential in teaching marginal analysis. We have argued that fudging

together with students' tendency to memorise concepts inevitably impedes their understanding of marginal analysis.

We strongly believe that the method we have suggested can be complementary with introductory microeconomics textbooks. Instructors of introductory microeconomics courses are responsible for supporting the textbook material with some subtlety that is missing from it – that is the value added of lectures after all – and hopefully this paper provides them with some guidance.

References

Gans, J, King S. P. & Mankiw, N. G. (1999) *Principles of Microeconomics*. Sydney: Hartcourt Brace.

Jones, C. (2005) *Applied Welfare Economics*. Oxford: Oxford University Press.

McTaggart, D., Findlay C. & Parkin, M. (2003) *Microeconomics*, 4th edition. Australia: Pearson Education.

Sloman, J. & Norris, K. (2005) *Principles of Microeconomics*. Australia: Pearson Education.

Swann, M. & McEachern, W. A. (2003) *Microeconomics*: A Contemporary Introduction, 2nd edition. Australia: Thomson.

Taylor, J. & Frost, L. (2002) *Microeconomics*, 2nd Australian edition. Milton: John Wiley and Sons.

Notes

The author would like to thank Phong Ngo, Ben Smith, Shane Evans and two anonymous referees for valuable comments. The usual disclaimer applies.

1 These include Gans *et al.* (1999), McTaggart *et al.* (2003), Sloman and Norris (2005), Swann and McEachern (2003), and Taylor and Frost (2002).

2 Of course, we are postulating that the price is above the minimum point of the *AV C* curve.

3 $MR > MC$ when $q = 12$ as this is a discrete example. The focus of our paper hereafter shall be on continuous cases.

4 Checking if Point *E* lies above the *AV C(q)* curve is the next step, but that is not the focus of our discussion.

5 For example, see Swann and McEachern (2003), p.225.

6 The motivation of introducing this rule is clear. This marginal analysis leads to a supply curve of a competitive firm that conforms to the law of supply – the supply increases as the price increases – which does make some students realise how powerful marginal analysis is.

7 In fact, if we assume that the total cost function is not too concave when the level of production is low, the corresponding *MC* curve is U-shaped and has a low vertical intercept. In that case, for a sufficiently high market price, we only have one intersection to worry about. Diagrams in McTaggart *et al.* (2003) and Sloman and Norris (2005) seem to be drawn in this spirit, but it is not clear whether they are conscious of this matter.

[8] Another typical fudging in introductory microeconomics is teaching consumer surplus as the measure of consumer's welfare. We all know that it is not an appropriate measure of consumer's welfare in general – see Jones (2005) for discussion – but to illustrate gains from trade in introductory microeconomics, it does the job.

[9] None of the textbooks referred to in this paper covers it.

[10] Of course, instructors need to derive the *MR* curve prior to this, which is important.

Contact

Akihito Asano
School of Economics (Building 25A)
Faculty of Economics and Commerce
The Australian National University
Canberra
ACT 0200
AUSTRALIA

Tel: +61 2 6125 3363
Fax: +61 2 6125 5124
Email: akihito.asano@anu.edu.au