

Transaction Cost and the Law of Demand

De-Xing Guan *

Abstract

In this paper we build a model with Coasian transaction cost to derive the demand curve without explicitly using the concept of utility. We show that when transaction cost is decreased due to technological change, improved institutions, among others, the goods demanded will be increased, since their expenditures are lower and producers of these goods will supply more or better goods to the market. Given utility or pleasure in consuming goods, a cost-based demand function would be derived and is negatively sloped. The law of demand can therefore be based only on cost, but this cost must consist of both prime cost and transaction cost. Without transaction cost there would be no market, and therefore there would be no price. Law of demand is thus more closely related to transaction cost than to utility.

Keywords: Law of Demand, Transaction Cost, Adam Smith, Ronald Coase

JEL classification: A1, D1, D2

* Department of Economics, National Taipei University, 151, University Road, San-Xia, New Taipei 23741, Taiwan (email: guan@mail.ntpu.edu.tw).

I. Introduction

Law of demand is one of the most important theorems in economics. It is usually represented by the demand curve derived from consumer's utility maximization problem. Adam Smith knew the concept of utility, but when the law of demand (and that of supply) first appeared in the *Wealth of Nations*, he did not use utility to explain how demand and supply work together to obtain the market price. His price theory was based on *value in exchange* rather than *value in use*, the former being related to cost and the latter to utility.¹ Using utility to derive demand function was not the idea of classical economists such as Smith and Ricardo, but that of Jevons, Menger, and Walras in the 1870s.² In this paper we would like to derive a cost-based theory of the law of demand. This does not mean that the utility-based demand theory is wrong. Rather, it is still a correct theory of demand if we admit the existence of the concept of utility. What we want to do in this paper is to propose a cost-based theory of demand such that even if utility does not exist or cannot be properly measured in reality, the law of demand is still well-defined. In this sense this paper is a complement, not a substitute, of the utility-based theory of demand.

In addition to Smith and Ricardo, Coase was also skeptic about the idea of utility, as he said that “[F]or group of human beings, in almost all circumstances, a higher (relative) price for anything will lead to a reduction in the amount demanded. This does not only refer to a money price but to price in its widest sense.” (1988, p. 4) and “[P]rice theory...does not seem to me to require us to assume that men are rational utility maximizers” (p. 5).³ According to these passages, it seems that he tried to base the law of demand not on utility but on cost. We share Coase's idea in this paper.

But to derive a cost-based theory of the law of demand we have to incorporate Coasian transaction cost into it. As argued by Smith, consumer's choice is indeed a make-or-buy decision based on the comparison of costs.⁴ For example, when people want to have a cup of coffee, aside from particular preferences, they can make it by

¹ As he said: “The word VALUE...sometimes expresses the utility of some particular object, and sometimes the power of purchasing other goods which the possession of that object conveys. The one may be called ‘value in use;’ the other, ‘value in exchange.’” (1776, Bk. I, Ch. IV) Friedman made it clearer: “The classical writers (Smith, Ricardo, etc.)...arrived at a labor cost theory of value, wherein utility was regarded as a condition or prerequisite of value but not as a measure of it.” (1976, p. 36)

² For a history of the development of utility theory, see Stigler (1950).

³ Coase once described utility as “a nonexistent entity which plays a part similar, I suspect, to that of ether in the old physics.” (1988, p.2)

⁴ Smith said: “It is the maxim of every prudent master of a family never to attempt to make at home what it will cost him more to make than to buy.” (1776, Bk. IV, Ch. II)

themselves or just buy one in the nearby café. The choice would be the one with least cost. Assume that this cost can be separated into two parts: prime cost and transaction cost. Prime cost consists of the original resources needed to produce the good. Transaction costs, as argued by Coase (1988), include at least costs of searching for information, bargaining and negotiation, and enforcing contracts. The cost-based approach we adopt here is to use the full cost, that is, prime cost plus transaction cost, to derive the law of demand, which is consistent with the cost theory of value pioneered by Smith and Ricardo.

One example can be used to explain why transaction cost is crucial to deriving the demand curve in which utility is not explicitly used. Assume that the prime cost of a cup of coffee is one dollar. This consists of the minimum cost of such factors of production as coffee bean, labor, and coffee machine, among others, in which the coffee could not be made if the producer would not pay this minimum cost of production. But the market price of coffee would be higher than the prime cost. This is because there are transaction costs, such as costs of retailing and wholesaling, say Z dollars, to bring coffee to the market. The price of coffee should be $1+Z$ dollars in a competitive market because these firms could only earn zero economic profits.⁵ Without these intermediate firms consumers should spend W more dollars to make it by themselves, and the full cost of this home-made coffee is $1+W$ dollars.⁶ This implies that you should do all things the café would do by yourself: to buy beans and coffee machine in the market, and to hire someone in the labor market or employ yourself to make the coffee.

The make-or-buy decision is to compare the full cost of making and that of buying, or between $1+W$ and $1+Z$ dollars. Consumers would choose to make coffee by themselves if $1+W < 1+Z$, or simply $W < Z$, and to buy one in the market if $W > Z$. And if $W = Z$, then it makes no difference how you get the coffee. Assume that those people who buy coffee in the market are less efficient, or having a lower productivity, in making coffee. Perhaps this is because they do not have time, place, machine, or skill to make coffee themselves. On the other hand, those who make coffee themselves would in general have higher productivity in making coffee because they might have a better way doing this. Note that when there are no transaction costs ($W=Z=0$), as assumed in the standard model of perfectly competitive equilibrium, the decisions of making and buying coffee are always equivalent because both full costs are equal to

⁵ What these firms of retailing and wholesaling earn are accounting profits (equal to Z dollars).

⁶ There are, of courses, other ways to have a cup of coffee. For example, you can buy a cup of coffee at other stores or even vending machines. To avoid unnecessary complication without changing the basic arguments in this paper, we simply focus on the make-or-buy example as described above.

the one-dollar prime cost, or $1=1$! This implies that autarky is as efficient as competitive markets are.

As emphasized by Coase (1988), markets exist because they can facilitate the transaction between buyers and sellers, bring about division of labor and specialization, and reduce the cost of production. But using markets is costly. If it were costless, then the market itself would no longer exist. People in such a world would provide themselves with everything they want without cooperating and transacting with each other. This is a world of autarky where demand is always equal to supply with the price of demand being the same with that of supply for all levels of output, and both the demand and supply prices would be equal to the prime cost of production because there are no transaction costs. In autarky both demand and supply curves are horizontal and overlapping with each other.⁷ On the contrary, when there are transaction costs the demand and supply curves could be separated. Markets would emerge to facilitate transactions if the cost is less than the benefit of using them. The law of demand, together with the corresponding negatively sloped demand curve, is therefore closely related to transaction costs in a cost-based theory of demand.

The paper is organized in the following way. Section II presents the benchmark model. Section III illustrates how to use Coasian transaction costs to derive demand curves without explicitly using the idea of utility. Section IV concludes.

II. A Model with Coasian Transaction Costs

Friedman (1976, p. 16) had defined that “The demand curve...represents the maximum quantities that would be purchased at alternative prices.”, and the demand curve could be generated by having “[T]he conditions affecting demand unaltered and at the same time to have a maximum of change with respect to the forces affecting supply.” (p. 30) In fact, consistent with the argument of Friedman, the empirical or statistical demand curve is usually generated by shifting supply curves to identify the demand curve. The cost-based theory of demand would do the same thing, where the factors affecting supply curves come from changes of the transaction cost in production. For example, when transaction cost is reduced through, say, more information, better negotiation and bargaining, or more smoothing legal process in

⁷ This is actually Keynes’s view of *Say’s Law*: “[S]upply creates its own demand in the sense that the aggregate demand price is equal to the aggregate supply price for all levels of output and employment.” (1936, pp. 21-22) What Keynes ignored was that the condition for his view to prevail is to assume zero transaction costs.

enforcing contracts, the product of the good would increase and the supply curve shifts to the right. The intersections of the quantities of the good produced and the various supply curves would thus trace out a demand curve. In the following we show how this could be done in a simple model of demand with transaction costs.

Assume that consumers face a Smithian make-or-buy decision: to make the good by themselves or to buy it in the market. The purpose of consumers is assumed to get the good they want in the least costly way. According to the principle of comparative advantage, sellers in the market are usually better at producing goods than buyers. Because using markets is costly, buyers should pay transaction costs such that sellers are willing to bring goods to the market. The cost minimization problem of consumers can be described as follows:

$$(1) C = Y \min(P_E, w^{1-\alpha} r^\alpha)$$

where C is total cost of producing the final good, say Y cups of coffee, P_E is the price of a unit of A_E which represents the idea or expertise a typical professional coffee maker in the market would have in making coffee, and $w^{1-\alpha} r^\alpha$ is the unit cost of the ordinary people in making coffee at home using their labor and physical capital as inputs, where $0 < \alpha < 1$.⁸ Professional coffee makers use the expertise (at the cost of $P_E A_E$) together with their labor and capital (at the expertise-denominated cost of $P_E Y - P_E A_E$, where Y is denominated in expertise) to produce the final product. The full cost for the professional coffee maker is the sum of transaction cost and prime cost, or $P_E A_E + (P_E Y - P_E A_E) = P_E Y$. The full cost for the home coffee maker is $w^{1-\alpha} r^\alpha Y$. Equation (1) indicates that ordinary consumers could either produce coffee for themselves, or buy it in the market. They just choose the least costly way to have a cup of coffee.

In equilibrium the cost of making or buying coffee would be the same. Because there are three inputs: professional coffee maker's expertise, labor, and capital, the total cost function can be written as $C = P_E A_E + w L_N + r K_N$, where w is wage rate, r is rental price of capital, and L_N , K_N are labor and capital of the ordinary people, respectively.⁹ We assume that both labor and capital markets are competitive but the market for expertise is not. The solution for this cost minimization problem requires

⁸ The expertise of professional coffee makers could be their knowledge concerning coffee, their skills in making coffee, or any other know-how which ordinary people could not easily obtain.

⁹ As will be shown later, the expertise is in turn produced by both labor and capital, and the aggregate production function will be a weighted average of the outputs produced by people at home and those in the market, with the weights being the fractions of labor and capital allocated to these two kinds of production.

that the total cost of making and buying coffee would be the same in equilibrium, so we have

$$(2) \quad C = P_E Y = w^{1-\alpha} r^\alpha Y$$

This means that $P_E Y = C = P_E A_E + wL_N + rK_N$, or $Y = A_E + (wL_N + rK_N)/P_E$. Since the unit cost function is assumed to be Cobb-Douglas, an immediate implication of this result is that $(1-\alpha)(wL_N + rK_N) = wL_N$. Combining this and the above equations we have

$$(3) \quad Y = A_E + (wL_N)/[(1-\alpha)(w^{1-\alpha} r^\alpha)]$$

By Shephard's Lemma: $L_N = \partial C / \partial w = (1-\alpha)w^{-\alpha} r^\alpha Y$, $K_N = \partial C / \partial r = \alpha w^{1-\alpha} r^{\alpha-1} Y$, so $K_N / L_N = \alpha w / [(1-\alpha)r]$. Inserting this into (3) and rearranging terms would have

$$(4) \quad Y = A_E + A_N K_N^\alpha L_N^{1-\alpha}$$

where $A_N = (1/(1-\alpha))((1-\alpha)/\alpha)^\alpha$.

To introduce Coasian transaction cost into our model let us assume that some efforts X are necessary in using markets to produce goods. These efforts include, for examples, searching for information, bargaining and negotiating, enforcing the contracts, and measuring the quality and quantity of goods. Without loss of generality, assume that the efforts of using markets are linearly related to the expertise of professional coffee makers, or assume that $A_E = \mu X$, where $\mu > 0$ is a variable representing the efficiency of using efforts to produce the expertise. A larger value of μ implies that professional coffee makers have better expertise such as more information, better knowledge and know-how, better skills in making coffee, and so on. Let the price of efforts be P_X , that is, the cost of a unit of efforts in terms of the final good. Note that $1/\mu$ is the cost of producing a unit of expertise in terms of efforts, so P_X / μ is the cost of producing a unit of market-made goods in terms of the final good, which we define as marginal transaction cost (C^T) of producing market-made goods, or $C^T = P_X / \mu$. The efforts of using markets are factors of production and therefore intermediate goods of producing final goods. They are themselves produced by other factors of production such as labor and capital. Assume that this production function is Cobb-Douglas: $X = K_E^\beta L_E^{1-\beta}$, such that we have

$A_E = \mu X = \mu K_E^\beta L_E^{1-\beta}$, where L_E , K_E are labor and capital devoted to the accumulation of expertise, respectively, $0 < \beta < 1$. We assume $\beta > \alpha$. This means that the marginal productivity at the extensive margin (acquiring expertise) is greater than that at the intensive one (no-expertise efforts), or that the production function of goods made in the market has larger marginal product than that made at home. Otherwise there are no consumers who would buy goods in the market if their qualities or convenience are the same. The above Cobb-Douglas function, together with equation (4), gives rise to the following aggregate production function:

$$(5) Y = \mu K_E^\beta L_E^{1-\beta} + A_N K_N^\alpha L_N^{1-\alpha}$$

Before the completion of the model, we first explore the relation between the aggregate production function and the market equilibrium of final goods. Suppose that there is a firm which wants to enter into the extensive margin (say, to start a café). The firm has two options: produce the market-made goods by itself or buy them from other firms in the market. If it chooses the former it becomes a seller of the good, and it becomes a buyer if it chooses the latter. According to Coase (1988), transaction costs are the costs involved in using institutions such as markets, firms, and the law. Now let the sum of discounted profits of acquiring new expertise be $R^T A_E$ such that R^T is the unit-rent of the market-made good. After producing final goods the firm has to provide some efforts for using the market. The price of doing this is P_X , and the price of buying the final goods from other firms in the market is P_E , so the firm would become a seller of the expertise if $P_E A_E \geq R^T A_E - P_X X$, and it would be a buyer of it if $R^T A_E - P_X X \geq P_E A_E$. And in arbitrage equilibrium it must be that $R^T A_E - P_X X = P_E A_E$. Note that $P_X = C^T \mu$ and $A_E = \mu X$, so $P_X X = C^T A_E$. This implies that $R^T A_E - C^T A_E = P_E A_E$, or simply $R^T - C^T = P_E$, or

$$(6) P_E + C^T = R^T$$

Equation (6) differs from the usual market equilibrium condition in two aspects. First, when there are no transaction costs ($C^T = 0$), $P_E = R^T$, and this is the standard arbitrage equilibrium condition: at the margin, the cost of buying the good is equal to the discounted sum of profits (or monopoly rent) generated by selling this good in the market. But when $C^T \rightarrow 0$, $\mu = P_X / C^T \rightarrow \infty$, so $X = A_E / \mu \rightarrow 0$: no efforts will be devoted to using the market. This contradicts the fact that using markets is costly in the real world. The second aspect is that when a firm would like to buy goods in the

market place it must pay the costs involved in using the market. If it does pay the full price, that is, prime costs plus transaction costs, then its demand for the good becomes Smith's *effectual demand*; otherwise, it is an *absolute demand*.¹⁰ Obviously here the effectual demand is represented by the full price $P_E + C^T$ such that without paying for transaction costs, the firm's demand would become absolute and it will not be realized in the market. The firm must pay not only the prime cost but the transaction cost to bring the good to the market. The firm would buy nothing if it only pays for the fixed cost. Another implication of equation (6) is that, for any goods to be effectively brought to the market, marginal benefits (rent) must exceed marginal costs (transaction cost) of doing so, or $R^T > C^T$. If the benefit fails to be larger than the cost, no new goods would be created. In the extreme case that $C^T \rightarrow \infty$, it is too costly for the firm to start a new business, such that there are no new goods to be produced at all. Mathematically, $A_E = P_X X / C^T \rightarrow 0$ as $C^T \rightarrow \infty$.

To close this model we need market-clearing conditions for both labor and capital. Assume that there is a θ fraction of people who would like to work at the extensive margin, and the remaining $1 - \theta$ at the intensive one, where $0 < \theta < 1$. For simplicity, we also assume that the proportions of capital employed at these two margins are the same as those of labor. Again nothing important would be changed if this assumption were relaxed. The labor and capital markets clear if $L_E + L_N = L$ and $K_E + K_N = K$, where L, K are the aggregate supply of labor and capital, respectively. When all markets clear, equation (5) would become

$$(7) Y = \theta \mu K^\beta L^{1-\beta} + (1 - \theta) A_N K^\alpha L^{1-\alpha}$$

Equation (7) characterizes the aggregate production possibility frontier. It is a weighted average of the production functions at extensive and intensive margins. All of the above results can be illustrated by Figure 1. In a world without transaction costs, no market-made goods would be produced because using the market is not costless. This implies that $A_E = 0$, and the point B in Figure 1 will shrink to the origin immediately. In a world with positive transaction costs there are two situations. First, if transaction costs are no less than the rent the firm might earn from its production of the new good, that is, if $R^T \leq C^T$, then obviously no goods will be produced. The point B in Figure 1 will again shrink to the origin. Second, if $R^T > C^T$, then the new good will be produced, and in equilibrium, $R^T - C^T = P_E > 0$, a positive price which is necessary for A_E to exist.

¹⁰ For the distinction between effectual demand and absolute demand, see Smith (1776, Bk. I, Ch. VII).

Transaction costs, therefore, act as thresholds to the introduction of new ideas or new goods into the economy. When transaction costs are lower because of better legal system, more information, less unnecessary lawsuits, less political conflicts, among others, point *B* in Figure 1 will move rightward to point *B'*, and the intersection point of the two production functions (point *A*) will move upward along the production curve at the extensive margin to another *newer* extensive margin (to point *A'* in equilibrium). This is because now the firm would have better expertise due to the reduction of transaction costs. This process will go on and on if more transaction costs are reduced and therefore better institutions are established. The long-run aggregate production possibility frontier will be the *upper envelope* of the production functions at various margins. There is always another better extensive margin out there for people to pursue if they can find a better way to get to it.

III. Law of Demand without Utility

In the first chapter of his *Principles*, Ricardo (1817) made the following statement: “Utility then is not the measure of exchangeable value, although it is absolutely essential to it...Possessing utility, commodities derive their exchangeable value from two sources: from their scarcity, and from the quantity of labor required to obtain them.” This is in fact a re-statement of Smith’s view on the distinction between value in use and value in exchange, as discussed in Section I. In this Section, based on the benchmark model in Section II, we would like to derive the demand curve without explicitly using the idea of utility. This does not mean that utility has nothing to do with the demand function. We just follow the arguments of Smith and Ricardo to assume that people might or might not have utility toward commodities, but the exchangeable value (or price) of them would be measured in costs. In other words, we derive demand curve not from utility maximization, but from cost minimization.

We use Figure 1 to explain how to derive the cost-based demand curve. In the upper part of this Figure there are three production functions: P1, P2, and P3. Since cost function is defined as the inverse of production function, of which the marginal products of various factors of production (expertise, labor, and capital) are assumed to be decreasing, marginal cost curves are therefore increasing and are defined as supply curves, as illustrated by S1, S2, and S3 in the lower part of the Figure. The production function P1 represents the technology a person would have if he/she decides to make coffee at home. Because in general there is no special expertise required to do this,

his/her marginal productivity is assumed to be lower than that of those people who make coffee and sell it in the market. The coffee maker in the market is therefore assumed to have a steeper production function as shown by P2. The curve S1 is the corresponding marginal cost curve, as well as the supply curve, to the production function P1. S1 is above the curve S2, which corresponds to the production function P2, because lower marginal productivity implies higher marginal cost.

Coasian transaction costs can be represented by the vertical distance between the supply curves S1 and S2, as indicated by CT1 in Figure 1, and this is why there are horizontal intercepts in the upper part of this Figure (B and B' not at the origin). The length of these intercepts is another way to measure transaction costs. When there are no transaction costs, there are no intercepts. For example, the length of the intercept between the origin and point B is measured by $P_E A_E$ which in equilibrium is equal to $(R^T - C^T)A_E$, so if transaction costs could be reduced the horizontal intercept would be lengthened to, say, point B'. This in turn implies that there is a technological progress or an institutional improvement such that the coffee maker in the market would have a better production function as represented by the curve P3 and a corresponding lower marginal cost curve indicated by the curve S3. Either higher marginal productivity or lower marginal cost means that the coffee maker could bring more and better coffee to the market, or that the supply curve would shift to the right from S2 to S3.

Now we are ready to show how the cost-based demand curve could be derived. In the original Smithian make-or-buy decision, the equilibrium would be reached at the intersection of the curves P1 and P2, or at point A in Figure 1. This determines the equilibrium quantity of expertise and the price of it is determined by the intersection of this quantity and the supply curve S1 because price is equal to full cost (prime cost plus transaction cost) in equilibrium and the S's curves consist of full costs. When transaction costs are reduced, the production function shifts from P2 to P3, and the corresponding supply curves shifts from S2 to S3. Again, the equilibrium quantity as determined in the upper part of the Figure would help determine the equilibrium price of the new expertise. As mentioned by Friedman (1976), demand curves could be identified by the shifts of supply curves, so if we trace out the intersections of equilibrium quantities of the expertise and the corresponding supply curves, then we have a negatively sloped demand curve by joining these intersections, as indicated by the curve D in the lower part of Figure 1. Nowhere the idea of utility is needed in the derivation of the demand curve except that there is a prerequisite that there are people who want to drink coffee.

IV. Concluding Remarks

Smith opposed utility in the *Wealth of Nations* because it is not utility (or value in use) but cost which is the relevant measure of price (or value in exchange). Ricardo shared the same idea with Smith. This Smith-Ricardo classical doctrine was entirely reversed in the 1870s because of the marginal utility revolution pioneered by Jevons, Menger, and Walras. But, after all, utility is a nonexistent entity, so we cannot use it to measure other existent entities. The measuring rod of value, especially the value in exchange, should be an existent entity, not just an imaginary concept such as utility.¹¹ In this sense the present paper is following the Smith-Ricardo-Coase tradition that it is cost, not utility, which is the proper measure of price, and this cost must consist of both prime cost and transaction cost. Without transaction cost there would be no market, and therefore there would be no price.

In this paper we use a simple model with Coasian transaction cost to derive the cost-based demand function without explicitly using the concept of utility. A decline of transaction cost due to technological progress, improved institution, among others, might reduce the expenditure cost of consumers, and they will certainly demand more goods. Given the utility (or *pleasure*) people might get from consuming goods, the law of demand is nothing more than saying that when the expenditure cost is lower, the goods demanded will be higher.¹² We do not need to know what the magnitude of pleasure the consumer would have, or what the marginal utility would be. Actually we know almost nothing about these. This is the law of demand we have ever been taught since the time of Smith and Ricardo.

¹¹ Smith opposed utility even in the moral sense. In the *Theory of Moral Sentiments* Smith criticized his best friend David Hume for basing moral judgments mainly on utility (1759, Pt. IV, Ch. I). But in this criticism, utility was rejected not because it is nonexistent, but because it is *sympathy* (as well as the *impartial spectator*), not utility, which is the main source of our moral sentiments.

¹² This is consistent with Coase's claim that "Why a man will take a risk of being killed in order to obtain a sandwich is hidden from us even though we know that, if the risk is increased sufficiently, he will forego seeking that pleasure." (1988, p. 5)

References

- Coase, Ronald H. (1988): *The Firm, the Market, and the Law*, Chicago: University of Chicago Press.
- Friedman, Milton (1976): *Price Theory*, 2nd edition, Chicago: Aldine.
- Keynes, John Maynard (1936): *The General Theory of Employment, Interest, and Money*, London: Macmillan.
- Ricardo, David (1817): *Principles of Political Economy and Taxation*, London: John Murray.
- Smith, Adam (1759): *The Theory of Moral Sentiments*, London: A. Millar.
- (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*, London: W. Strahan and T. Cadell.
- Stigler, George (1950): "The Development of Utility Theory. I," *Journal of Political Economy*, 58, 307-327.

Figure 1: Transaction Cost and Demand Curve

