The Transactions Demand for Cash: A Critical Analysis of Baumol’s Model

1. Introduction

The theory of the demand for cash based on the micro-economic theory of an optimal enterprise inventory was formulated by W.J. Baumol in his work *The Transactions Demand for Cash: An Inventory Theoretic Approach* written in 1952 [2, pp. 545–556]. It quickly challenged the then existing theories, explaining the principles of establishing cash balances by the market participants to maintain an undisturbed process of consumption. The theory contradicted commonly accepted approaches which stipulated those balances to be a more or less stable part of currently consumed incomes (see [3, p. 65]).

Baumol’s model was based on assets similar to money (in terms of their properties) earning interest as their income, and encumbered with a relatively small risk. Those assets gave economic entities a choice of keeping their income in cash or as interest-earning assets. Keeping cash balance involves – as any other choice made by the market participants – bearing certain economic costs. It is a natural tendency of all economic entities which are guided by rational premises either to bring those costs to a minimum or, alternatively, to maximize total utility derived from a certain solution. In his model Baumol accepted for the optimization criterion the search for such a volume of cash balance which would minimize the total cost of its maintenance. Hence, he presented a cost-based approach.

In his paper *The Interest-Elasticity of Transactions Demand for Cash*, published in 1956, J. Tobin (the winner of the Nobel prize in economics) presented a different, revenue-based approach to the demand for cash, formulated in terms
of the theory of inventories [6, pp. 241–247]. Despite – as it seems – its much higher precision in comparison to the cost-based approach, the revenue-based approach was not, regrettably, as popular as Baumol’s model. The basic form of the latter has been presented by the most currently published papers on the demand for cash (see, e.g. [1, pp. 125–133], [3, pp. 65–67], [4, pp. 50–56], [5, pp. 416–419]). This is surprising, especially in view of the fact that Tobin’s analysis indicates that Baumol’s function of economic costs related to cash balances held by the economic entities sets them lower than the real balances. In his review of Baumol’s model (attached as an enclosure to this paper) Tobin points out that problem. However, he does not come up with an alternative cost-based approach (see [6, p. 247]). This paper presents an attempt at verification of Baumol’s model in this respect, and was written after its author considered such an approach substantiated.

2. Baumol’s model

As already mentioned, Baumol’s model of the demand for cash based on the theory of inventories rests on the principle of minimizing economic costs by the entities participating in the market. That model constitutes, which should be openly stated, nothing but an explanation of the transactional premise for demanding cash. Baumol assumed a priori, to free himself from prudent and speculative motive of maintaining cash balances by some economic entities, that all market transactions are perfectly predictable, and all economic entities’ consumer expenditure is uniformly distributed in time (which means, that the same volume of expenditure is allocated to the same time unit).

To simplify the analysis, apart from the above provisions, Baumol made one more implicite assumption about the uniformity of interest rate in the scale of the whole economy. This meant in particular the same market interest rate both for the deposits (regardless of their character) as well as loans. It should be added that Baumol understood savings – also implicite – as any form of keeping assets over time, except for keeping cash, i.e. the assets with a zero interest rate. An additional simplification of the model was the fact that it corresponded only to static economy. The model completely neglects the problem of money changing its value in time (which the author chose not to mention), the consequence of impact of the monetary factor represented by changing prices within economy.

Baumol analyses three various scenarios for an economic entity striving to define its optimum cash balance necessary to conduct operations on the market (i.e. a balance allowing undisturbed consumption and, at the same time, minimizing the costs):
1) an economic entity finances consumption exclusively from its earlier accumulated savings;
2) an entity finances consumption exclusively from external sources, i.e. cash credit;
3) an entity finances consumption from its own resources. However, these are not savings, but currently generated revenues.

The first two instances may be considered jointly from the angle of the costs of transactions. This is feasible only due to the provision enforcing uniform interest rate for the whole economy, which implies the same economic cost of keeping cash balances, regardless of the source of transaction money (cash). This is because the cost of lost interest due to the withdrawal of a certain amount of interest-earning assets (vide case one) will equal the interest charged on credit money (vide case two).

Baumol expressed those cases with the use of models in the following way:

An entity needs a certain $T$ amount of money to provide for consumer expenditure within a certain unit of time. The entity finances that expenditure either through withdrawal of the whole (or a part) of its deposits, or through debt, assuming that in the beginning it does not have to contribute cash balance amounting to $T$, but it can draw cash in equal $C$ installments (equal both in time units, and in terms of their amount).

Cash balances involve, however, two general categories of costs:
– lost opportunity costs (or costs incurred to pay back credit debt) of interest calculated with the $i$ market interest rate, and
– broadly understood transaction costs – conversion of interest-earning assets into cash $b$ (or the cost of credit), which include not only the cost of alternative time devoted to those operations (the so-called “costs of going about”), but also commissions charged by the financial institutions on such operations, which Baumol considers as fixed costs$^1$.

With such formulation of the problem, and keeping in mind the earlier assumptions, an entity must set up $C$ cash balances within a certain time limit (in terms of their number). This involves total transaction costs equal $b \frac{T}{C}$. Since

---

$^1$ Actually – Baumol makes that point very clear – the costs of transactions should be described with the use of $b+kC$ function. That is because they are often divided into a part independent of the value of the transaction of converting interest-earning assets into cash and vice versa ($b$), and the part directly dependent on that value (i.e. $kC$). However, Baumol notes that the value of those costs might be considered lump-sum (the same for each transaction), hence no serious implications for $b+kC$ function are caused by their introduction into the model. See [2, p. 546 (footnote 2) and p. 547 (footnote 4)].
the consumer costs are uniformly distributed in time, an average cash balance owned by a certain entity equals \( \frac{C}{2} \) which, according to Baumol, allows to calculate the cost of servicing the balance within the whole time in question as \( i \frac{C}{2} \). Hence, it is obvious that the total cost of a cash balance to an entity is the sum of costs of setting it up and maintaining, thus:

\[
b \frac{T}{C} + i \frac{C}{2}.
\]  

(1)

The key problem in Baumol’s model is the search for the already defined optimal cash balance which needs to be set up by an entity to finance its consumer needs. The formula deciding its volume is reached by calculating the derivative of the function of total economic costs expressed with the formula (1) to calculate \( C \). Hence, the derivative:

\[
-b \frac{T}{C^2} + \frac{i}{C} = 0,
\]  

(2)

after transformation yields the following formula allowing to calculate the sought for volume of \( C \):

\[
C = \sqrt{\frac{2bT}{i}}.
\]  

(3)

In this way the basic assumptions of the discussed model are reached. The above derived formula (termed in professional writings on the subject the square root formula – see [4, p. 53]) indicates, that a rational entity’s demand for cash is proportionate to the square root of the total value of transactions within a certain time, and inversely proportional to the square root of market interest rate. Additionally, it is suggested that the increase of transaction costs raises the cash balance of an entity to the extent equal to the square root of change of those costs.

The third case which Baumol considered is different from the discussed above since an entity has a possibility to transform a part (or the whole) of its cash revenue into interest-earning assets. It is obvious that a rational entity will invest all revenues exceeding the planned current expenditures – naturally, unless the transaction costs (transforming cash into interest-earning assets and vice versa) are higher than generated revenues. This case requires supplementing the above model with the transformation costs involved in transforming a part (or the whole) of cash revenue into interest-earning assets.

With the use of the previously employed symbols this case may be described in a model form as follows:
At the beginning of a certain period an entity generates revenue equal to the planned consumer expenditure, i.e. $T$ volume. It decides to transform $I$ volume into interest-earning assets, and leave the remaining part of revenue amounting to $R$ as a transaction money reserve. In view of the foregoing, the entity will be forced to establish again cash balance to finance consequent consumer expenditures after $\frac{T - I}{T}$ period. Since an average volume of cash in that part of period equals $\frac{T - I}{2}$, the cost of lost interest due to maintaining cash balance amounting to $R = T - I$ and making investment equal $I$ will be expressed with the following formula$^2$:

$$\frac{T - I}{T} \cdot \frac{T - I}{2} + b.$$

(4)

When the whole cash balance (the part of revenue that has not been invested) of a given entity has been spent, and the entity has not generated an additional revenue, then the necessity arises to liquidate the whole (or a part thereof) volume of investment (possibly with the accrued interest). In such circumstances an entity is faced with a decision dilemma described in the first case, the only difference being that it does not comprise the whole period, but its part defined as $\frac{I}{T}$. Hence, the total economic cost of transaction cash for that period will amount to:

$$\frac{I}{T} \cdot \frac{C}{2} + b \cdot \frac{I}{C},$$

(5)

which after differentiating relative to $C$ produces an identical with the two other cases formula defining the optimum volume of cash balance for that part of a period (that is after the initial volume of transaction cash has been fully spent) – vide formula (3).

Nevertheless, the crucial issue of that model situation is the definition of the optimum initial volume of an entity’s transaction cash. It is arrived at – similarly to all other cases – through the calculation of a derivative of the function of total economic costs of keeping transaction balances for the whole period, this time

---

$^2$ Even though in this case Baumol provides the analysis with the functions of transaction costs $b_d + k_d$ (for $d$ investment) and $b_w + k_w$ (for each withdrawal of $w$ interest-earning assets), this paper, similarly to preceding cases – for reasons described in footnote 1 – assumes lump-sum volume of these costs (the same for every transaction). Additionally, the same volume of costs was assumed, irrespective of the volume of transaction, i.e. no matter whether an entity is investing, or being liquidated (in fact, this is what Baumol decided himself – see [2, p. 548]).
relative to the volume of investment – a variable $I$. Hence, by differentiating the function:

$$\frac{T - I}{T} \cdot \frac{T - I}{2} + b + \frac{I}{T} \cdot \frac{C}{2} + b \frac{I}{C},$$

its derivative is as follows:

$$-\frac{T - I}{T} \cdot \frac{iC}{2T} + \frac{b}{C},$$

which after equation to zero and solution $R = T - I$ gives:

$$R = \frac{C}{2} + \frac{bT}{iC}.$$  

By transforming this formula to $R = \frac{C^2}{2C} + \frac{bT}{iC}$, a then substituting for $C^2$ in compliance with the formula calculated earlier (3) $\frac{2bT}{i}$, $R = \frac{2bT}{iC}$, is reached, which after the reversed substitution gives:

$$R = C.$$  

In this way the same conclusion is drawn in this case regarding the principles of maintaining the optimum transaction balance by an entity as in the two earlier variants.

### 3. The function of economic costs of Baumol’s model: an alternative approach

The above presented model of demand for cash, formulated by Baumol, does not raise any doubts regarding the correctness of its assumptions as to the principles of creating the demand for cash by the individual market entities, which is a prerequisite for their transactions. However, some serious reservations may be made about the principles on which the author structured his alternative costs of keeping cash (we mean the lost opportunity costs or interest charged on credit), particularly in the context of the function of transaction costs he derived.

As it can be seen from the model, Baumol defined the function of alternative costs of keeping transaction money as the volume of interest that was lost (or paid), calculated on the basis of an average volume of cash which an entity had
at a certain time. Hence, he completely overlooked the fact that an entity resigning from keeping a certain deposit as interest-earning assets (or going into debt), loses in this way revenue which would be earned from interest (or it is charged interest) on its total, and not just average, cash balance. For the sake of exercise, taking an extreme case when at the beginning of a period an entity establishes transaction balance equal to the total value of transactions within that period, it may be concluded in conformity to Baumol’s reasoning that the volume of lost (or paid) interest is calculated exactly on the half of the transaction’s value, and not on the whole (as it should be). It is clear, that this whole sum lowers the value of interest-earning assets or denotes its credit debt.

Baumol’s reasoning might be theoretically substantiated (at least partially) only under the circumstances, when an entity conducted an infinite number of opening cash balance transactions to maximally reduce the cost of lost (or paid) interest. Such a solution would involve – due to the infinite number of transactions – huge transaction costs, whose volume would also grow infinitely. However, even under such circumstances – owing to the fact that Baumol formulating his function of alternative costs of possessing transaction cash completely neglected the problem of the moment, and the number of transactions of establishing cash balances within a certain time – the absolute value of the error of measurement of the volume of those costs will not be subject to change.

In view of the foregoing Tobin’s theorem seems to be substantiated. He said, “Baumol’s calculation of the costs of interest is rather difficult to comprehend” [6, p. 247]. This is because sheer logic suggests that the volume of revenue generated by an entity from earned interest (or the cost paid for using money from credit) are equal to market interest rate at a certain period of time and an average balance of interest-earning assets (or an average balance of a credit account). This cannot be, by any means, reconciled with the definition of the problem contained by Baumol’s work. To make our point, let us prove the following theorem, maintaining all the earlier assumptions.

We will begin with the scrutiny of the first of cases analysed by Baumol – i.e. the case where an entity finances its consumption only from its earlier accumulated own savings.

To simplify the theorem let us assume that at the beginning of each period the entity’s savings equal its planned expenditures. Having a specified number of transactions of establishing entity’s cash balance within a certain period denoted by $\frac{T}{C}$, cash balance of interest-earning assets denoted by a will after each $k$ transaction in time will be denoted by an arithmetic sequence $(a_k)$ of a general term $a_k = T - kC$, where $k = 1, ..., \frac{T}{C}$. Calculating the sum of such defined arith-
metric sequence \[ \sum_{k=1}^{T} a_k = \left( \frac{a_1 + a_T}{2} \right) \frac{T}{C}, \] and then dividing it by the number of conversions of interest-earning assets into transaction cash we reach the formula which defines an average balance of an entity within the analysed period of time as \( \frac{T - C}{2} \). With the application of market interest rate we reach another formula denoting the volume of entity’s revenue (expressed in percentage points) which it absorbs from the volume of interest-earning assets earmarked for financing consumption within a certain period of time, but which are gradually liquidated, i.e. \( i \frac{T - C}{2} \). It is obvious that had that entity maintained its savings intact for the whole analysed period of time, it would have earned interest equal \( iT \). This leads us to the conclusion that the cost of lost interest, which is the difference between its potential volume and the volume actually earned, will be denoted in this case with the following formula:

\[ \frac{i(T + C)}{2}. \]  

Considering the second case, when an entity finances its consumption from external sources (using credit money), then its debt \( b \) after each \( k \) operation of opening cash balance, their number denoted by \( \frac{T}{C} \), will be expressed by an arithmetic sequence \( (b_k) \), of a general term \( b_k = kC \), where \( k = 1, ..., \frac{T}{C} \). The quotient of the sum of this sequence \( \sum_{k=1}^{T} b_k = \left( \frac{b_1 + b_T}{2} \right) \frac{T}{C} \) and the number of times an entity goes into debt within that period of time denotes an average balance of its credit account within that time as \( \frac{T + C}{2} \), which after the application of market interest rate allows to calculate the cost of using credit money with the formula:

\[ \frac{i(T + C)}{2}. \]
Considering the fact that formula (10), which allows to calculate the cost of lost interest (vide case one), is identical with formula (11), which denotes the volume of interest charged to the user of credit money (vide case two), the function of total economic costs of maintaining cash balances by an entity covering both cases may be defined as:

\[ C_T \beta C_T i \pm \frac{T}{2} + \frac{T}{C}. \]  \hspace{1cm} (12)

Comparing it with formula (1) it is clear to see, that Baumol’s function does not include the interest corresponding to an average sum of entity’s consumer expenditure within a certain period of time, i.e. \( i \frac{T}{2} \).

We still have case three to consider, that is the situation when an entity generating revenue equal to its consumer expenditure planned for that period decides to keep a part of revenue in cash, and invest the remaining part, i.e. convert it into interest-earning assets.

When that entity starts that period with an \( R \) transaction balance, which is the difference between the total of consumer expenditures planned for that period denoted by \( T \) and the investment denoted by \( I \), it is obvious that it will lose revenue (in the form of interest) from the whole sum, equal to market interest rate charged for the whole period, that is \( ii(T - I) \). The formula defining the volume of lost interest for the remaining time (i.e. after the initial volume of transaction money has been spent) may be constructed in the same way as for case one. Since we know that the potential interest to be earned within that time frame, which is \( \frac{I}{T} \) of the whole period, on the whole investment outlay amounts to \( \frac{I}{T} iI \), and the real interest earned by the entity on its assets in that time equals \( \frac{I}{T} i \frac{I - C}{2} \), as proved above, then the volume of interest lost during that period of time will be expressed by \( \frac{I}{T} i \frac{I + C}{2} \).

Adding up all cost components borne by an entity in the third case we arrive at the function of total economic costs related to holding transaction money resources as:

\[ i(T - I) + b + \frac{I}{T} i \frac{I + C}{2} + b \frac{I}{C}. \]  \hspace{1cm} (13)
Again, comparing it with the formula (6) which is proper for that case, we come to the conclusion that Baumol’s function underestimates the volume of costs, this time the volume being \( \frac{T - I}{2} + \frac{I}{T} \).

4. Conclusion

The above presented considerations allow to prove an a priori Baumol’s theorem of the function of total economic costs due to holding by market entities cash balances erroneous. Although it has no consequences on final implications of the model (the first derivatives of functions that were corrected in this paper yield the same result as Baumol’s functions), it still has to be stressed that each economic model must display internal integrity. Hence, it is to be hoped that the authors of papers on microeconomic theory of demand for cash in terms of the theory of inventory will bring more attention to that matter, this way reducing the number of misled readers. Maybe it will be worth their while to become acquainted with the above quoted J. Tobin’s work which, apart from W.J. Baumol’s work, they keep presenting as the foundation of their theories?

References


About the Author

Artur Pollok engages in research and teaching at the Cracow University of Economics. He graduated from the University in 1996 with an M.A. In 2000 the Research Council of the Faculty of Economics approved continuation of his PhD research thesis on “Household Savings in Poland during the Transformation Period of the Economic System”. His PhD research is financed by the State Committee for Scientific Research in
Warsaw in the form of a grant. At present he is a teaching assistant in the Microeconomics Department of the Faculty of Economics of the Cracow University of Economics. He is a permanent member of the Polish Economic Society and – since 1997 – has been a member of the Presidium of the Society’s National Board in Warsaw. He also acts as Board Secretary for the Cracow Branch. His research mainly concentrates on the following: microeconomics, transformation of the economic system in Poland, household economics, savings and consumer processes, social aspects of economic management.

e-mail: pollok@pte.pl