

鲍莫尔模型和惠伦模型述评

盛 松 成

凯恩斯在《通论》中,将人们对货币的需求分为交易动机的货币需求、预防动机的货币需求和投机动机的货币需求三种类型,并认为,出于前两种动机的货币需求依存于收入的多少,即 $M_1 = L_1(Y)$,而出于后一种动机的货币需求则依存于利率的高低,即 $M_2 = L_2(r)$ 。于是,凯恩斯的货币需求函数就成了: $M = M_1 + M_2 = L_1(Y) + L_2(r)$ 。五十年代后,西方经济学家纷纷著文,丰富和发展了凯恩斯的货币需求理论,其中比较有影响的是:鲍莫尔(William Baumol)等人对凯恩斯关于交易性货币需求理论的发展、惠伦(Edward Whalen)等人对凯恩斯关于预防性货币需求理论的发展和托宾等人对凯恩斯关于投机性货币需求理论的发展。因本文篇幅所限,同时还由于托宾等人的理论同另一重要的货币需求理论——资产选择理论有着不可分割的联系,所以,拟将它放在笔者即将发表的《资产选择理论的演变》一文中介绍,而本文只评析鲍莫尔模型和惠伦模型。^①

鲍莫尔模型

早在四十年代末,美国著名经济学家汉森(A. H. Hansen)就对凯恩斯关于交易性货币需求主要取决于收入的多少,而同利率的高低基本无关的观点提出了质疑。他指出,当利率上升到相当的高度时,货币的交易余额也会具有利率弹性。^②五十年代初,美国经济学家鲍莫尔第一次深入分析了交易性货币需求与利率的关系。^③以后,托宾也论证了货币的交易需求同样受到利率变动的影响。^④这就是为西方经济学界所广泛接受的鲍莫尔—托宾模型。^⑤

鲍莫尔模型的基本思路是这样的:“人们收入的获得和收入的使用一般不会同时发生,且可假设支出的发生是逐渐的、平稳的”^⑥,因此,人们没有必要让所有的收入都以现金形式存在(因为现金不会给其持有者带来收益),而可以把其中的大部分暂时贷放出去,以获取利息,以后再逐渐收回贷款,以供日常支付之用。当然,将非现金资产变为

现金,需要一定的手续费,鲍莫尔把它称为“佣金”(brokerage fee)。但只要利息超过手续费,这样做就是有利可图的。而利率越高,

-
- ① 在这两个模型中,货币都只限于现金,故本文以下所说的货币也只是指现金,两词通用。
 - ② 汉森:《货币理论与财政政策》,英文版,1949年,第66—67页。
 - ③ 鲍莫尔:《现金的交易需求——一种存货的理论分析》,载《经济学季刊》1952年11月。在他的《经济理论和经营分析》(英文版,1977年)一书中,鲍莫尔又概括了他在上文中所提出的观点。
 - ④ 托宾:《现金的交易需求的利率弹性》,载《经济学和统计学评论》1956年8月,收入《经济学论文集》(英文版,1971年)第一卷。
 - ⑤ 鲍莫尔模型和托宾模型的基本结论和分析方法在本质上并无二致,但“前者由于把货币需求问题同存货理论(inventory theory)联系起来了,因此更令人感兴趣”(H. G. 约翰逊:《货币经济学论文选》,英文版,1978年,第27页),且前者比较简单明了,故本文只介绍鲍莫尔模型。读者若有兴趣了解托宾模型,请参阅拙作《流动性偏好理论的新发展》,载《金融研究》1985年第4期。
 - ⑥ 鲍莫尔:《现金的交易需求——一种存货的理论分析》,载《经济学季刊》1952年11月,第547页。

利息超过手续费的机会就越多,由非现金资产变为现金的次数也就可以越多,从而平均保留在身边的现金就会越少。可见,“如果企业都是有效的利润极大化者,凯恩斯贬低利率对现金的交易需求的影响就很可能是错误的。”^①“即使持有现金只是为了满足当前交易的需要,对现金的需求也同样是利率的反函数。”^②

为了具体展开分析,鲍莫尔还作了如下规定:

(1) 人们有规律地每隔一段时间获得一定量的收入(Y),而他们的支出则是连续的和均匀的。

(2) 放款一律采取购买政府短期债券的形式,因为这种贷款最安全。

(3) 每次出售债券与前一次出售的时间间隔及每次的出售量(K)都相等。

鲍莫尔认为,“一个企业的现金余额通常可以看作是一种存货(inventory)——一种货币的存货,这种存货能被其持有者随时用来交换劳动、原料等。这种存货同鞋子制造商准备随时用以交换现金的鞋子的存货并没有本质的区别。”^③保存任何存货都需耗费成本。保存鞋子的成本是保管费用和占用资金所须支付的利息等。那么,保存现金的成本又是什么呢?鲍莫尔指出,保存现金的成本有两项。一项是获取现金(即出售债券)时所须支付的手续费(b)。根据上述第三条规定,每次债券出售额都为 K ,而支出总额则为 Y ,故在整个支出期间内,全部手续费为 $b(Y/K)$ 。还有一项就是持有现金所牺牲的利息(以 r 表示利率)。根据上述第一条规定,支出是一不变的流量,所以在整个支出期间的平均交易余额为 $K/2$,从而利息成本就为 $r(K/2)$ 。若以 c 代表保存现金的总成本,则有下式:

$$c = \frac{bY}{K} + \frac{rK}{2} \quad (1)$$

我们知道,任何企业只要不妨碍生产的

正常进行,总是力图将保持存货的成本降至最低。同样,鲍莫尔认为,现金持有者也都是有理性的,他们也务必获得保存现金的最低成本。为求得使 c 变为极小的 K 的值,我们对式(1)求 K 的一阶导数,并令其为零:

$$\frac{\partial c}{\partial K} = -\frac{bY}{K^2} + \frac{r}{2} = 0 \quad (2)$$

于是有:

$$K = \sqrt{\frac{2bY}{r}} \quad (3)$$

同时还可求得:

$$\frac{\partial^2 c}{\partial^2 K} = \frac{2bY}{K^3} > 0 \quad (4)$$

从而满足了式(3)的 K 使式(1)变为极小的必要条件。

如上所述,人们在整个支出期间内的平均交易余额为 $K/2$,所以,若以实际现金余额的形式来表示,使保持现金的成本变得最小的平均交易余额应为:

$$\frac{M}{P} = \frac{K}{2} = \frac{1}{2} \sqrt{\frac{2bY}{r}} \quad (5)$$

或

$$M = \alpha Y^{0.5} r^{-0.5} P \quad (5')$$

其中, P 为一般物价水平, $\alpha = \frac{1}{2} \sqrt{2b}$ 。

上式就是著名的“平方根公式”(square root formula)。它表明,“当交易量或佣金增加时,最适度的现金(存货)余额就将增加,而当利率上升时,这一现金余额就会下降。但存货分析表明,这种变化又不是成比例的。例如,最适度的现金余额只能作为交易量的平方根,随着后者的上升而上升——即在企业的最适度现金余额的决定中,有着规模经济的规律在起作用。”^④同样,虽然现金的交易需求将随着利率的变动而作相反方向的变化,但前者的变化幅度却比后者来得小。从

① 鲍莫尔:《经济理论和经营分析》,第493页。

② 托宾:《现金的交易需求的利率弹性》,见《经济学论文集》第一卷,第229页。

③ 鲍莫尔:《经济理论和经营分析》,第491页。

④ 同上书,第492页。

式(5')看,就是现金的交易需求的收入弹性和利率弹性分别为0.5和-0.5(绝对值都小于1)。

鲍莫尔模型自五十年代产生以来,始终受到西方经济学界的重视。六十年代后西方出版的有关货币理论和货币政策的著作几乎都对它有所论及。究其原因,当然是由于它对西方货币需求理论的发展和西方国家货币政策的制定产生了相当的影响。这些影响似可归纳为以下几点:

第一,鲍莫尔模型论证了最基本的货币需求——交易性货币需求也在很大程度上受到利率变动的影响。这一论证不仅为凯恩斯主义的以利率作为货币政策的传导机制的理论进一步提供了证明,而且向货币政策的制定者们指出,货币政策如果不能影响利率,那它的作用就会很小。

第二,根据平方根公式,假定利率和物价不变,增加一定比例的货币存量,将导致收入的更高比例的增加。因为货币的交易需求与收入的平方根成正比例变化,所以,为了使公众能吸收这部分追加的货币,收入的增加的比例就必须大于货币供给的增加了的比例。鲍莫尔模型由此强调了货币政策的重要性。

第三,鲍莫尔模型不仅阐明了收入的变化与交易性货币需求变化的数量关系,而且间接地说明了物价的变化与交易性货币需求变化的关系(因为物价上升使名义交易量和手续费都增加了。鲍莫尔并且证明,由于名义交易量和手续费的增加,物价的上升将使最适度交易性货币余额以同比例上升)。这对于确定最适度的货币供应量和保持货币市场的均衡,具有一定的参考价值。

其中最重要的是第一点。可以说,凯恩斯关于货币的交易需求并不受制于利率的论断,是以人们将把本期购买商品所需要的全部现金都保存在身边作为假设条件的,但他却没能证明这一假设的合理性。而鲍莫尔则从消费者并不需要持有整个交易期间将花费

的所有现金,而可以把暂时闲置的部分用来产生利息这样一个显而易见的事实出发,得出了与凯恩斯大相径庭的结论。就它的这一基本结论而言,鲍莫尔模型作为一种货币需求理论,具有一般性的意义。

当然,模型对利率变动及交易量变动与交易性货币需求变动的数量关系不一定描述得准确。西方经济学家对鲍莫尔模型的批评和修正大多也是围绕着这些数量关系进行的。如美国经济学家布伦纳(Karl Brunner)和梅尔泽(Allan Meltzer)认为,货币需求对交易量的弹性并不是一个常数,而是一个变数。当交易量变小时,这一弹性向下限0.5靠近;而当交易量变大时,这一弹性则向上限1靠近。^①美国经济学家米勒(Merton H. Miller)和奥尔(Daniel Orr)也根据他们的计量模型指出,“现金需求和交易量的联系是松散的,现金需求对交易量的弹性并不存在一个精确的值。”^②它可以在1/3至2/3之间,甚至在更大的范围内变化。鲍莫尔模型将这一弹性一律定为1/2,是它的“最明显的弱点之一”。^③还有人计算出,在资本主义国家中,交易性货币需求的利率弹性总是明显低于0.5。

造成鲍莫尔模型的上述失误的主要原因恐怕在于:第一,模型假设任何企业和个人的收支都是已知的、可预料的和完全有规律的。可这并不符合实际情况。意料之外的收入超出支出或支出超出收入是生活中的常事。前者往往使现金余额迅速增加,后者则很可能使现金余额急剧减少。所以,平均现金余额并不是很稳定的,交易性货币需求对交易量的弹性也并非总是0.5。

第二,模型假设企业只能通过借款或通过出卖金融资产获取现金,而不能通过出卖

① 布伦纳和梅尔泽:《现金余额中的规模经济的再思考》,载《经济学季刊》1967年8月,第427页。

②③ 米勒和奥尔:《企业货币需求的一个模型》,载《经济学季刊》1966年8月,第425、426页。

产品来获得。无论从整个社会看,或从个别企业的较长一段时间看,这一假设都是很不现实的。布伦纳和梅尔泽曾通过计量分析指出,鲍莫尔模型若摒弃这一假设,就不会得出现金的交易需求具有规模经济的性质这一结论。^①

第三,模型没能区分交易性现金需求对于不同收入的家庭和不同规模的企业所具有的不同的利率弹性。很难想象,一个年薪二十万美元的公司小职员在考虑如何处置近期内就要花费掉的小额现金时,同年薪二十万美元的公司总经理在斟酌以何种形式持有收入会给自己带来最大的利益时,会同样重视利率的因素,对于利率的变动会有同样的敏感性。又何以断言说,交易性现金需求的利率弹性总是 -0.5 呢?

由上可见,鲍莫尔模型对影响交易性现金需求的诸因素所作的定量分析具有很大的局限性,是不足为训的。

惠伦模型

凯恩斯不仅认为交易动机的货币需求与利率的高低无关,他还认为,预防动机的货币需求也不受利率变动的影响。鲍莫尔—托宾模型已经证明了凯恩斯上述第一个论点不能成立。那么,凯恩斯的第二个论点是否还能成为西方经济学界所普遍接受呢?回答也是否定的。1966年,美国经济学家惠伦及米勒和奥尔先后发表文章,论证了预防动机的货币需求也同样为利率的反函数。^②其中比较有代表性的是惠伦模型。

惠伦首先给预防性货币需求下了定义。他说,预防性货币需求来自事物的不确定性。一个人无法保证他在某一时期的货币收入和货币支出同事前预料的完全一致,也不可能排除实际支出超过实际收入或发生不测之事以至临时需要现金的可能性,因此,人们实际保持的货币往往比预期所需要的多一些。其中超额部分即来自预防动机的货币需求。

惠伦接着分析了影响最适度预防性货币需求的因素。他认为主要有三个因素,一是非流动性的成本(cost of illiquidity),二是持有预防性现金余额的机会成本,三是收入和支出的平均值和变化的状况。非流动性的成本是指“因低估在某一支付期间内的现金需要而造成的后果的严重性”,^③即少持有或不持有预防性货币余额的机会成本。如果人们不能得到足够的贷款,且又缺乏能迅速转换成现金的资产,当因缺乏现金而无法应付必需的支付时,他们就会陷于经济困境或遭致破产。在这种情况下,非流动性的成本是很高的;如果人们能够得到贷款,则非流动性的成本就取决于贷款利率的高低;如果人们持有较易转换成现金的资产(如各种票据),此时,非流动性的成本又等于由非现金资产转换成现金的手续费(类似于鲍莫尔模型中的佣金)。惠伦认为,企业和个人不能保证他们能随时获得所需要的贷款,故对于具有高度理性的人们来说,这第三种情况应作为理论分析的一般对象。

所谓持有预防性现金余额的机会成本则是指为持有这些现金而须放弃的一定的利息收益(它又类似于鲍莫尔模型中的持有交易性现金余额的机会成本)。上述非流动性的成本和持有预防性现金余额的机会成本构成了持有预防性现金余额的总成本。

现在,货币持有者面临着一个难以两全的局面:“如果他为预防不测持有较多的货币,他就减少了预期非流动性的成本,但同时却增加了持有预防性现金余额的机会成本”;反之,如果他持有较少的预防性现金余额,他

① 见布伦纳和梅尔泽:《现金余额中的规模经济的再思考》。

② 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月;米勒和奥尔:《企业货币需求的一个模型》,载《经济学季刊》,1966年8月。

③ 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月,第315页。

就减少了持有预防性现金余额的机会成本,可又提高了非流动性的成本。所以,“欲使利润极大化的企业的目标就是选择某一最适度的预防性现金余额,以使这两种成本之和下降到最低限度。”^①惠伦随后要讨论的就是如何使这一总成本降至最低。

若以 r 代表利率, M 代表预防性现金的平均持有额,则持有预防性现金余额的机会成本就为 $M \cdot r$ 。若以 c 代表每次将非现金资产转换成现金的手续费(为了分析的简便起见,假设 c 的大小不受转换量的影响), P 代表一定期间内这种转换的可能的次数,则预期非流动性的成本就为 $P \cdot c$ 。若以 E 代表总成本,可有下式:

$$E = M \cdot r + P \cdot c \quad (6)$$

需要进一步探讨的是, P 是如何决定的。我们知道,只有当一定期间内支出和收入的差额〔惠伦称之为净支出(net disbursements)〕大于该期间内预防性现金的持有额时,才需要将非现金资产换成现金。为了求得出净支出超过预防性现金余额的可能性(亦即出现的概率),惠伦引用了特奇庇切夫(Tchebycheff)的不等式定理(inequality theorem)。根据此不等式定理,一个变量偏离其中位数(偏离度为其标准差与乘数 K 的乘积)的概率等于或小于 $1/K^2$ 。在惠伦模型中,这一变量就是净支出。由于假设人们的长期目标是要花费掉所有的收入,故净支出的概率分布趋于零,亦即净支出的中位数为零。于是所求概率实际上就是偏离零的概率。若以 s 代表净支出分布的标准差, K 代表它的一个乘数, M 代表预防性现金余额,就有下式:

$$M = K \cdot s \quad (7)$$

上式表示,支出与收入之间的差额(即净支出偏离其中位数的值)需由预防性现金余额来弥补。

根据特奇庇切夫的不等式定理,净支出超过预防性现金余额的概率可由下式表示:

$$P \leq \frac{1}{K^2} \quad (8)$$

将式(7)代入式(8),得:

$$P \leq \frac{1}{(M/s)^2} \quad (9)$$

惠伦假设,企业和家庭都是风险回避者,所以,“在估计净支出超过预防性现金余额的可能性时,最保守的估计应是对预防性现金管理的最适当的估计”。^②式(9)的不等式于是变为以下等式:

$$P = s^2/M^2 \quad (10)$$

将式(10)代入式(6),持有预防性现金余额的总成本就成了:

$$E = M \cdot r + (s^2/M^2)c \quad (11)$$

对上式求 M 的一阶导数,并令其为零:

$$\frac{\partial E}{\partial M} = r - \frac{2s^2}{M^3}c = 0 \quad (12)$$

从而有:

$$M = \sqrt{\frac{2s^2c}{r}} \quad (13)$$

并可求得式(11)对 M 的二阶导数大于零:

$$\frac{\partial^2 E}{\partial^2 M} = \frac{6s^2}{M^4}c > 0 \quad (14)$$

可见,由式(13)所确定的预防性现金余额 M 将使持有该余额的总成本 E 降至最低限度。

式(13)表明,“最适度预防性现金余额的变化分别同以下三个变量的变化成立方根的关系:(1)净支出分布的方差 s^2 , (2)非流动性的成本 c 和(3)持有现金余额的机会成本率 r 的倒数。”^③式(13)由此证明了前述惠伦关于影响最适度预防性现金需求的三个因素的论断,尤其是证明了预防动机的货币需求同样受利率变动的影响。利率越高,此项货币需求越小;利率越低,此项货币需求越大。

① 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月,第318页。

② 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月,第318页。

③ 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月,第319页。

预防性货币需求的利率弹性为 $-1/3$ 。需要解释的是第三个因素——收入和支出的影响是如何在式(13)中表现出来的。

式(13)表明,预防性货币需求将随着净支出分布的方差(为标准差之平方) s^2 的立方根的变化而变化。因此,根据式(13),预防性货币需求同收入和支出的数量关系,取决于净支出分布的状况。因收入和支出的数额和次数都将影响净支出的概率分布,故两者的变化都会引起预防性货币需求的变化。“如果每次收入和支出的平均额保持不变,而收入和支出的次数变了,净支出的正态分布的标准差将随着收入和支出额的平方根的变化而变化,也就是这一分布的方差将与收入和支出额以相同的比例变化。所以,根据上述最适度预防性现金余额的公式,当假定一种净支出的正态分布后,最适度预防性现金余额将随着收入和支出的平均额的立方根的变化而变化。”^①也就是预防性现金余额的需求对收入和支出的弹性为 $1/3$ 。“另一方面,如果收入和支出的次数保持不变,而它们的数额增加了,这一方差将随着收入和支出额的平方的变化而变化,由此而确定的预防性现金需求的收入和支出弹性为 $2/3$ 。”^②

同鲍莫尔模型一样,惠伦模型的基本结论——预防性货币需求也受利率变动的影响,

同样反映了资本主义经济中的一般情况,但惠伦模型对影响预防性货币需求的诸因素所作的定量分析,又是值得商榷的。因篇幅所限,本文对此不作进一步讨论。

根据鲍莫尔模型和惠伦模型,凯恩斯的货币需求函数应被修正为 $M = L_1(r, Y) + L_2(r)$,并可进一步简化成 $M/P = L(r, Y)$ 。这也就是凯恩斯学派的货币需求函数。它表明,对实际货币余额的需求是由利率和收入两个因素共同决定的。

现代西方经济学家所提出的理论一般都是为他们的政策建议服务的,鲍莫尔模型和惠伦模型当然也不例外。但这两个模型本身是属于技术性的,且我国居民的货币需求大致也可分为交易性货币需求和预防性货币需求两种类型(我国基本上不存在投机动机的货币需求)。因此,在我们确定流通中的货币必要量,研究利率和收入与居民持币量以及与消费和储蓄的关系时,在我们研究如何利用利率杠杆来实行国家的货币政策时,这两个模型也可作为我们的参考。

① 惠伦:《现金的预防需求的合理化》,载《经济学季刊》1966年5月,第318、319页。

② 理查德·科格伦(Richard Coghlan):《货币金融理论》,英文版,1980年,第92页。

(上接第30页)

的历史环境分辨斯大林经济思想中的正确与谬误。总之,要把斯大林的经济思想放在社会主义实践史和马克思主义经济学说史这个大的历史背景下来加以研究。可以这样说,对斯大林经济思想作出正确的分析和评价,就能比较容易地说明五、六十年代以后各社会主义国家经济理论发展变化的来龙去脉。因为各社会主义国家的经济体制改革都是针对在斯大林思想影响下建立起来的旧体制进行的,社会主义经济理论的突破性进展正是从批判斯大林“模式”开始的。

关于研究毛泽东经济思想的重要性,是大家所瞭解的,不需多说。

除了上面这些,需要研究的问题还有许多,例如,近几年新发表的马克思、恩格斯、列宁、毛泽东等人的著作、手稿需要进行深入的分析研究;卢森堡、布哈林、铁托和卡德尔等重要人物的经济思想有待进一步研究;关于一百多年来马克思主义经济学说同资产阶级经济学说的多次大论战的情况,需要进行分阶段的研究并在此基础上做系统的全局性的分析等。不过,相比而言,应该把更多的力量放在上述三个方面。