

货币增长与通货膨胀

在上一讲中，我们探讨了货币制度的基础框架及其在现代经济中的核心作用。本讲将在此基础上进一步理解货币政策制定的依据，通过四个紧密相连的部分揭示货币增长与通货膨胀之间的复杂关系。首先，我们将分析货币的需求与供给，探讨中央银行如何通过调控货币供应来应对经济情势变化。随后，通过利率模型，我们将了解利率如何作为经济活动与货币政策之间的桥梁，影响家庭与企业借贷、消费与投资决策。继而，本讲将引导大家探讨货币市场的均衡及货币数量论，理解货币供应量的变化是如何影响整体经济的价格水平和产出的。最后，通过对通货膨胀、反通胀及通货紧缩的成本的讨论，我们将分析这些经济现象对社会、企业和消费者的具体影响，以及政策制定者如何试图通过不同的策略来管理这些风险。通过串联这些概念，我们将全面理解货币政策的重要性以及它在维护经济稳定中的关键作用。本讲旨在帮助读者系统地理解货币政策的关键作用及其影响经济稳定的机制，揭示货币动态与经济波动之间的复杂联系。货币政策的重要性可以从长期和短期来说。从长期看，货币政策能维持币值稳定，从而使经济稳定；从短期看，货币政策影响GDP、就业、出口、进口甚至股市。我们将在之后的学习中更深入地讨论短期货币政策，而本讲更侧重长期的逻辑。

第一节 货币需求与供给

一、货币的短期需求

人们为什么需要持有货币？这个问题是理解货币需求的关键。在短期，人们对货币的需求主要来源于以下三个方面。

（一）货币的交易需求

货币的**交易需求**（**transaction demand for money**）是个体和企业为完成日常购买和销售活动而持有货币的需要，来源于货币作为交易媒介的基本功能，这是因为在现代经济中，货币是买卖商品和服务的普遍接受的支付方式。为了更好地理解该维度，我们不妨讨论一个稍微有年代的模型——鲍莫尔-托宾（**Baumol-Tobin**）模型。该经济学理论描述个体如何在持有货币和其它资产之间做决策，由威廉·鲍莫尔¹和

¹ Baumol, W. J. (1952). The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics*, 66(4), 545-556.

詹姆斯·托宾²在 1950 年代独立提出。其核心思想是：持有货币虽然方便交易，但也会带来机会成本——将资金以非货币形式持有能获得更多利息收益（利率）。该模型通过数学公式计算出了一个最优的货币持有量，以平衡交易成本（如将资金转换为货币时的手续费）和机会成本。理想情况下，个体会选择一个使总成本（交易成本和机会成本之和）最小的货币持有水平。该最优持有量与几个关键变量有关：个体的总交易额、转换成本以及可获得的替代资产的利率。这个模型在经济学中有着较重要的地位，因为它提供了一个基本的理论框架，用来理解和预测货币需求的变化，特别是在利率变动时对个体和整体经济活动的影响。

鲍莫尔-托宾模型的基本设定是：人们获得的收入 $\$Y$ 直接存入（名义）利率为 i 的银行账户（或购买名义利率为 i 的短期债券），但把这部分钱转化为货币 M 需要支付一笔固定费用 C ，并且这笔费用与涉及金额的大小无关——这就是取钱的单位交易成本。如果一段时间内取钱的频次为 N ，则总交易成本为 $N \times C$ 。假设人们在期末时恰好花完所有的收入（没有储蓄），且期间保持匀速的花钱速率。

基于以上信息，我们能看出持有货币的机会成本就是 $i \times M$ ，因为在该设定中，持有货币本身不带来利息，或者说持有其它资产和持有货币的利差就是 i 。货币平均持有量 M 等于多少呢？给定取钱频次 N ，我们不难算出： $M = \frac{\$Y}{2N}$ 。为什么呢？下图是一个例子。

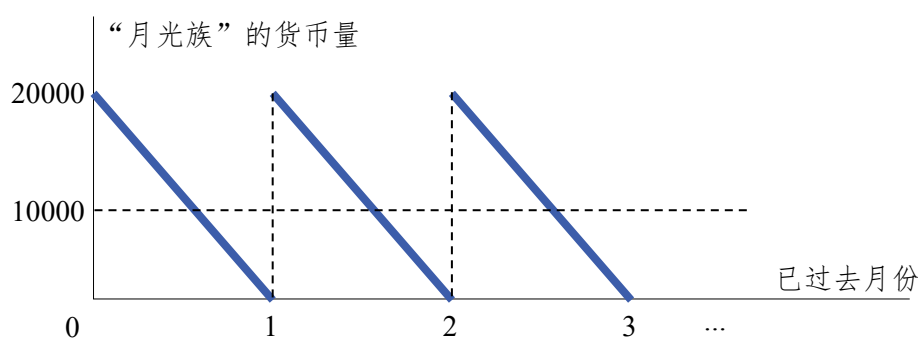


图 9.1：一个“月光族”的货币持有量动态

我们考虑一个“月光族”，他的年薪是 240000 元。如果他取钱的频率是每个月一次，那么一年期间内取钱频次就是 12 次。这样一来，他每个月初时的货币持有量就是 $240000/12=20000$ 元（如图 9.1 所示）。由于我们假设了人们保持匀速的花钱速率，因此该货币持有量在两次取钱之间会保持匀速下降至 0 元，在此期间平均的货币持有量就是 $20000/2=10000$ 元，而这就是图 9.1 中标记出来的水平虚线。不难看出，在这个例子中，货币平均持有量 $M = \frac{240000}{2 \times 12}$ ，其中 $\$Y = 240000$ ， $N = 12$ 。

² Tobin, J. (1956). The interest-elasticity of transactions demand for cash. *Review of Economics and Statistics*, 38(3), 241-247.

在模型中， N 不是外省给定的，而是内生决定的。如何内生化 N 呢？要回答这个问题，我们就要讨论经济学中最经典的问题——人们面临什么权衡取舍？根据基本设定，我们不难看出，取钱频率越高，总交易成本越高；但是，提高取钱频率可以降低平均机会成本，因为平均货币持有量 M 会随着 N 的提高而下降，那么这样收入在大部分时候会以产生利息的资产的形式存在。人们需要平衡这两种成本，使得两者之和最小化。

总成本就是总交易成本加上持有货币的（平均）机会成本。根据上面的计算，我们可以写出：

$$\text{总成本} = N \times C + i \times \frac{\$Y}{2N} \quad (9.1)$$

如何最小化该成本呢？该总成本函数是关于 N 的函数，如果我们画出总成本与 N 的关系曲线，会发现该曲线会先下降后上升。那么，在该曲线的“拐点”处，就有总成本的最小值。在这个“拐点”，曲线的斜率为0（即过该点的切线是水平线），那么在数学上就等价于函数对 N 的一阶导数为0：

$$\frac{\partial \text{总成本}}{\partial N} = C - i \times \frac{\$Y}{2N^2} = 0 \quad (9.2)$$

不难解出最优的取钱频次和平均货币持有量：

$$N^* = \left(\frac{i \times \$Y}{2C} \right)^{0.5} \quad (9.3)$$

$$M^* = \frac{\$Y}{2N^*} = \left(\frac{C \times \$Y}{2i} \right)^{0.5} \quad (9.4)$$

这样一来，我们就得到了交易需求派生出来的货币需求函数。这一函数与交易成本、总收入和利率有关：（1）交易成本越高，平均货币持有量越大，因为人们会减少取钱频率，从而导致手中平均持有更多货币；（2）总收入越高，平均货币持有量也越大，因为人们消费需求增大，需要更多货币用于交易；（3）利率越高，平均货币持有量越小，因为持有货币的机会成本增大。

（二）货币的谨慎需求

货币的**谨慎需求**（precautionary demand for money）是为了防止意外发生时产生临时花费或收入损失而持有货币，有时也被成为预防性需求。

谨慎需求的大小受到诸多因素的影响。第一，不确定性程度。当个体预期未来收入和支出存在较大不确定性时，谨慎需求要求持有更高的货币持有量。例如，当个人预期健康冲击将在一段时间内不定时出现时，会持有“备用”的货币，并且在两次取钱之间可能不会花这个“备用”的货币。这样导致的结果就是图 9.1 中曲线整体向上平移。第二，风险厌恶程度。当个体对不确定性的厌恶程度越高（即容忍程度越低）时，就会持有更多“准备”货币。第三，信贷条件。假如人们可以轻易超额花费，那么就能缓解和降低预防性需求。然而，当信贷市场紧缩时，获取贷款比较困难，人们可能还是需要持有一定的“备用”货币。第四，收入和财富水平。通常情况下，收入越高，个体承受不确定性的能力越强，这样看起来收入对谨慎需求和交易需求的影响是相反的。当然，若不确定性和风险厌恶程度是以比例而非绝对值表示的，当收入越高时，意味着交易需求越大，而不确定性和风险厌恶程度是交易需求的一定比例；若要求个体等比例持有“准备”货币，将会导致“准备”货币的绝对值更大。总之，收入和财富水平对谨慎需求的影响可能相对复杂一些，很难说是减少还是增加，但可能还是倾向于抵消部分由交易需求带来的货币持有量。³第五，利率。更高的利率提高了持有货币的机会成本，从而也抑制了谨慎需求，这一点与交易需求是一致的。

（三）货币的投机需求

货币的**投机需求**（speculative demand for money）是出于随时准备买入有价证券的考虑而持有货币的需要。该需求与债券等其它生息资产的真实利率有关。

假设市场真实利率为 r 。当 r 下降时，人们手中持有的老债券的价格上升（因为它比新发行债券能提供更高的利息），人们可能会因此出售债券而持有更多货币。我们在第五讲例 5.1 中也给出了一个利率下降导致老债券价格上升的例子，而这里则指出当债券价格上涨时人们“套现”的投机行为导致货币需求的增加。从另一个角度想，当 r 下降时，新债券的收益率不佳，人们可能会主动持有更多货币以待未来利率回升后再购入收益率更好的新债券。这么看来，投机需求引至的货币持有量与利率之间也存在负向关系，这与交易需求是一致的。

投机需求与交易需求不同的一个新视角是可以讨论**预期**的作用。若市场参与者预期 r 会下降（无论当下 r 是高是低），他们可能提前购买债券而持有更少货币，以等待未来利率降低时再卖出价格上涨的旧债券；反之，若市场参与者预期 r 会上升，他们可能主动持有更多货币，以待未来利率上升时能买入收益率更好的新债券。

拓展阅读 9.1

鲍莫尔-托宾模型虽然在经济理论中具有创新性和广泛的应用，但也存在一些限制和缺陷。

³ 从现实的角度来看，更高的收入和财富水平一般意味着更强的风险承受能力、更高的财务灵活性和访问其它金融工具的能力。这些都意味着拥有更高收入和财富水平的个体不需要因为谨慎动机而持有那么多货币，即使他们的交易需求可能更高。

首先，该模型的假设是理想化的。例如，交易成本和机会成本是恒定的，以及个人在每次转换中都转换相同金额的资金。现实中，这些假设可能并不总是成立。

其次，该模型忽略了风险因素，如持有现金可能会遭受贬值货被盗窃的风险。

第三，该模型是一个静态的模型，没有考虑时间跨度内的动态变化，如经济周期或者利率长期变动的影响。

第四，经济主体行为模式的简化。比如，该模型假设经济主体是完全理性的，总是能计算出并选择最优货币持有量。然而，现实中个体行为可能受到认知偏误和信息不完全等因素的影响。

第五，忽略技术变革的影响。随着电子支付技术和数字货币的发展，个人和企业对于实物现金的依赖程度降低，这可能改变货币需求的结构和动态。

尽管存在这些限制，鲍莫尔-托宾模型仍然是理解和分析货币需求特别是在不同利率环境下的一个重要工具。经济学者常常结合此模型与其他理论，以更全面地解释和预测经济行为。

例 9.1

在电子支付技术和数字货币发展的今天，还存在鲍莫尔-托宾模型中所说的交易成本吗？

现如今，鲍莫尔-托宾模型中讨论的传统意义上的交易成本（如将资金从银行账户转换为现金的成本）确实有所减少。现代支付系统提供了快速、低成本甚至成本为零的资金转移方式，这些方式显著降低了个人和企业在进行交易时的摩擦。然而，这并不意味着交易成本完全消失了。在现代经济中，虽然转换成本降低，但仍存在其他类型的成本如：

- **安全成本**：电子支付系统需防范黑客攻击和欺诈行为，这涉及到数据保护和安全性投资的成本。
- **隐私成本**：使用电子支付系统可能涉及个人数据的收集和处理，用户可能对隐私侵犯有所顾虑。
- **兼容性和可用性成本**：新兴支付技术需要与现有系统兼容，且用户需要技术知识和设备来接入这些服务。

经济学者已经在考虑如何将这些新兴技术整合到货币需求的理论框架中。例如，一些研究开始集中于数字货币和电子支付方式对货币政策、货币需求和总体经济活动的影响。这些研究考虑了技术创新带来的效率提升和成本降低，以及这些变化如何重新定义货币的需求和价值。此外，学术界和政策制定者也在关注如何通过改进的统计模型来更好地预测和分析在数字化、去现金化趋势下的经济活动。这包括考虑非传统数据源（如大数据和机器学习技术），以提高对货币政策影响的理解和预测能力。

上面讨论的货币需求的三个来源在凯恩斯 1936 年发表的《通论》中早有涉及。经典的凯恩斯货币需求理论提出了货币需求的三个动机——交易动机、谨慎动机和投机动机，它们解释了为什么个人和企业需要持有货币。凯恩斯强调了货币需求不

仅是为了交易目的，还涉及更广泛的经济行为，包括应对风险和市场机会。这三个动机是现代宏观经济学中分析货币市场的基础之一。

例 9.2

凯恩斯将货币需求称为“流动性偏好”，因此，我们一般讨论的货币需求量是在 M1 层面的，包括现金和活期存款，是人们想要花但还没花的货币量。基于此，你认为下面各项是我们一般（狭义）讨论的货币需求量吗？

- 人们放在银行的定期存款
- 人们已经花出去的钱
- 人们借给他人的钱

人们放在银行的定期存款属于 M2，不是我们狭义上讨论的货币需求量。它通常具有一定的存期，提前取款时可能会受到罚款或无法获得完整的利息收益，因此流动性较低，不像活期存款那样随时可用于日常交易。定期存款更多地被视为一种储蓄或投资工具，人们通过将资金存入定期账户来获得相较活期存款更高的利息收益。这表明，定期存款主要满足的是个人或企业的储蓄或投资需求，而非直接的交易需求。

人们已经花出去的钱也不是货币需求的一部分。货币需求通常指的是个体或企业在某一时间点上持有的货币量，包括现金和易于提取的存款，用于日常交易、预防不确定性支出或投机。这里关键是“持有”——即个体或企业选择在家中保留多少货币来满足各种需要。当钱被花出去时，它已经执行了其作为交易媒介的功能，不再是持有者的货币需求的一部分。这些资金转移到了另一个个体或企业的手中，可能会成为其货币需求的一部分，或继续在经济中流动。这些支出反映了货币的流通性，但不直接反映特定时间点上个体或企业的货币需求量。

人们借给他人的钱是否应该被视为货币需求的一部分，取决于我们选取的观察视角。当钱被借出时，它已经离开了原持有者的控制，因此不再计入其个人或企业的货币需求。当然，对于借款人而言，这些资金可能暂时成为其货币需求的一部分，尤其是当他们希望使用这些资金进行消费、投资或其他需要即时支付的活动时。在这种情况下，借入的钱就增加了借款人的流动性，可能也增加了他们的货币需求量。

从广义上讲，M1 只是显现出来的货币需求量，而还有一些货币需求量则是未被观察到的，这就包括立马可以从银行账户里变现出来的定期存款、立马可以卖掉的债券、立马可以卖掉的房子等等。这些未被显示出来的货币需求在人们真的需要时可以从 M2、M3 甚至 M4 转变为 M1，从而显现出来。

二、货币的长期需求

在长期中，货币需求的驱动力与短期内的一些日常波动不同，长期需求更多地受到结构性和基本经济变量的影响。短期货币需求受到即时经济状况和政策变动的强烈影响，反应迅速且直接。当我们扩展时间框架，考虑到长期因素时，我们必须关注更为持久的变量，如经济增长、人口结构变化、技术进步以及全球经济环境的

演变。这些因素虽然变化较慢，但对货币需求的影响深远且持久。例如，从长期的视角来看，物价水平是影响货币需求的核心因素。根据稍后要介绍的货币数量论，长期的货币需求与物价水平有着直接的正比关系。随着物价水平的上升，人们需要更多的货币来进行日常交易和维持购买力，反之亦然。此外，技术发展如数字支付和加密货币的兴起也正在改变我们对货币的需求和使用方式，这些变化在长期内可能重新定义货币的角色和功能。因此，虽然短期货币需求的变化对制定当下的货币政策至关重要，但了解和预测长期货币需求对于制定未来的经济策略同样不可或缺。这不仅帮助我们理解货币政策的长期效果，也为我们提供了评估经济健康的一个更广阔的视角。

我们这里主要讨论长期货币需求量与物价水平的关系。事实上，两者的关系并不难理解。当物价水平高了，人们就要多花钱，自然得多持有货币；反之，当物价水平低的时候，人们只需要持有更少的货币。

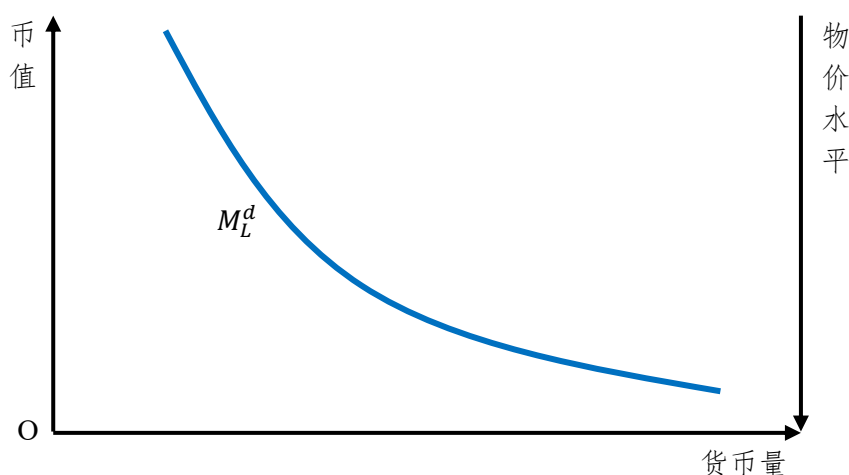


图 9.2：长期货币需求量与物价水平或币值的关系

有的人可能会问：为什么货币价值越高，人们愿意持有的货币反而更少了呢？这里存在一个误区，就是将持有货币当成财产保值的需要。实际上，财产保值这一任务应该交给其它金融或非金融资产。我们这儿讲持有货币的目的应该是为了花钱方便。持有货币是有机成本，其流动性高，但相比其它资产回报率低。若货币价值高了，那么不需要持有那么多就能达成交易、谨慎或投机的目的了，自然可以减少货币持有并增加其它资产。

在图 9.2 中，我们可以将物价水平记为 P ，那么单位币值就是 $1/P$ 。 M_L^d 的上标 d 代表需求，而下标 L 代表长期。为什么利率没有出现在这张图上呢？实际上，利率 r 只是被隐去了。在长期，真实利率不取决于货币供求，而是取决于可贷资金市场上真实投资与真实储蓄之间的关系。根据图 5.19（费雪效应），物价水平变化或通货膨胀只是改变名义利率，而不是真实利率。我们第二节会具体展开讨论。

三、货币需求函数

奥利弗·布兰查德（Olivier Blanchard）的经典教材《宏观经济学》中给出：

$$M^d = \$Y \times L(i) \quad (9.5)$$

其中， $\$Y$ 与前面鲍莫尔-托宾模型一样是指名义 GDP，而 i 是名义利率，而 $L(\cdot)$ 是一个减函数。布兰查德并没有强调这是长期货币需求函数，但其包含了长期和短期的考虑。注意，这里 $i \approx r + (\text{预期}) \text{通货膨胀率}$ 。

在公式 (9.5) 中，虽然物价水平并没有明着写出来，但是 $\$Y$ 里就包含了真实 GDP 和物价水平两个因素，可以写成 $P \times Y$ 。根据公式 (9.4)，我们得到 $L(\cdot)$ 的一个可能的表达式：

$$L(i) = \left(\frac{C}{2\$Y} \right)^{0.5} \times i^{-0.5} \quad (9.6)$$

1995 年诺贝尔经济学奖得主罗伯特·卢卡斯（Robert E. Lucas, Jr., 1937-2023）与合作者于 2021 年在经济学顶级期刊 *Journal of Monetary Economics* 上发表文章，提供了长期货币需求的证据。⁴他们获取了 38 个国家的数据进行实证研究，支持了长期稳定的 $L(\cdot)$ 函数的存在。在他们的研究中，货币需求就是对 M1，而名义利率则使用短期利率，比如 3 个月国库券的利率。

对于图 9.3 (a) 中的 12 个国家，证据非常有力，但有两个重要注意事项：对于巴西和阿根廷，有几个数据点明显低于理论线。在这两个案例中，这些点对应它们成功稳定恶性通胀之后的几年。图 9.3 (b) 中接下来的 14 个国家也清楚地显示出了负相关的证据，尽管在某些情况下（如巴林、巴巴多斯和泰国），斜率似乎与鲍莫尔-托宾模型中的线性技术（linear technology）⁵所给出的 0.5 有所不同。图 9.3 (c) 顶行的六个国家也提供了良好的证据，尽管在某些情况下（如芬兰和西德），情况并不像前面那么清晰。最后，图 9.3 (c) 底部行的六个国家问题最多。葡萄牙以及巴拉圭和挪威似乎也显示出负相关的证据，但与理论曲线相差甚远。从这些交叉图的角度来看，最后三个案例是彻底的失败。

⁴ Benati, L., Lucas Jr, R. E., Nicolini, J. P., & Weber, W. (2021). International evidence on long-run money demand. *Journal of Monetary Economics*, 117, 43-63.

⁵ 在经济学中，特别是在鲍莫尔-托宾模型中，线性技术是指生产函数或相关经济模型被设定为了线性形式，这意味着产出与投入之间存在直接且固定比例的关系——无论投入水平如何，边际产出保持不变。具体而言，在鲍莫尔-托宾模型中，总交易成本 $N \times C$ 与交易次数 N 呈现线性关系，机会成本 $i \times M$ 与货币持有量 M 也呈现线性关系。线性技术假设使得经济学模型容易计算，避免了非线性因素（如规模经济或递减边际回报等）的复杂性。虽然线性技术假设为模型提供了数学上的便利，但它并不总能准确地捕捉更为复杂的经济现象，如技术进步、创新对成本结构的影响，以及市场环境变化下的企业行为。

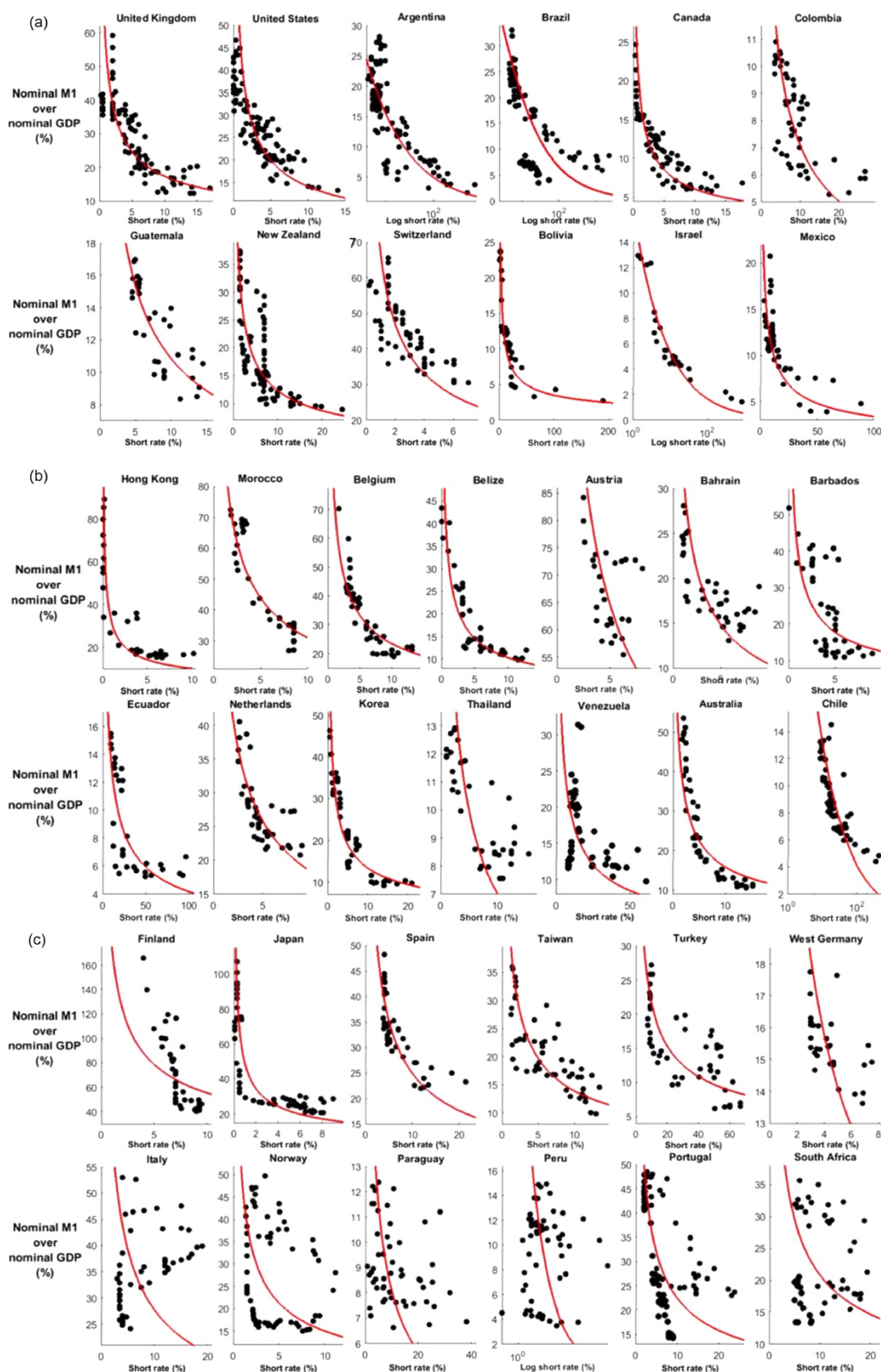


图 9.3: M1 与名义 GDP 之比和短期利率的关系

注：本图来自 Benati 等（2021）的论文。图中的散点是原始数据，而红色曲线则是基于理论绘制关系曲线（将利率弹性设置为 0.5）。

四、货币供给与长期供求平衡

在现实中，我们观察到的货币量也是货币供给量——它往往是由中央银行独立决定的、可控的，因此在长期不受物价水平（或币值）的左右，在短期也不受市场真实利率的影响。

我们之前已经讲过，中央银行通过公开市场操作、存款准备金率和贴现率调控银行系统重的货币供给量。与图 9.2 不同的是，即便中央银行通过改变利率来间接影响经济，货币供给曲线依然是垂直的。货币供给量的设定是基于中央银行的政策目标，如控制通货膨胀、支持经济增长等，而非市场利率的变化。因此，在模型中，货币供应量被视为政策决定的固定量，不随利率变化而改变。短期内，货币供给的独立性意味着其不直接依赖于市场利率的水平。虽然中央银行确实通过改变利率来影响经济活动和信贷创建，但这些活动反映在货币需求的变化上，而不是货币供给本身。在很多宏观经济模型中，将货币供给曲线设为垂直线是为了简化分析，因为这样可以清楚地区分货币政策的直接效应（通过货币供给控制）和间接效应（通过利率影响经济活动）。

同样地，货币供给量在长期的水平也不取决于物价本身，而是由中央银行根据其评估的经济需要独立设定，而不是由市场自发决定的。图 9.4 将货币长期供给与需求曲线绘制在了一个平面上，从而组成货币的长期供求模型。

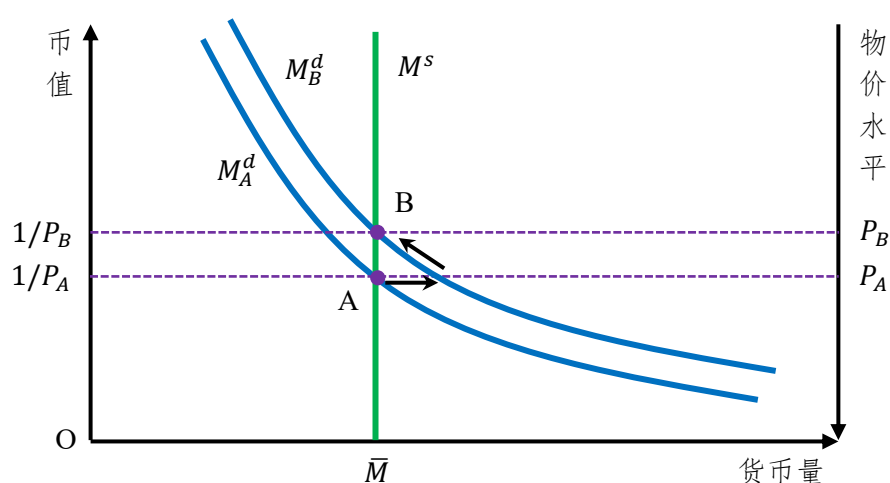


图 9.4：货币在长期的供求平衡

在这个模型中，均衡的偏离与恢复是依靠物价水平变化来实现的。假如货币的供给不变，但需求提高 ($M_A^d \rightarrow M_B^d$)，那么需求曲线就会向上移动。在 P_A 的物价水平上，货币的超额需求就会形成， $M_B^d(P_A) > M^s(P_A) = \bar{M}$ 。注意，这里 $M^d(\cdot)$ 和 $M^s(\cdot)$ 分别代表货币需求和供给函数。此时，人们会感觉手头紧、钱不够用，因此减少支出以增加货币持有，而这会使得物价下降，币值上升，直到超额需求不再出现为止，此时货币需求 $M_B^d(P_B) = M^s(P_B) = \bar{M}$ 。

第二节 利率模型

本节将讨论最基础的利率模型。在上一讲中，我们讨论了中央银行的货币政策工具可以改变货币供给量。接下来的问题是，改变货币供给量如何影响利率？利率究竟是怎么决定的？调控利率有用吗？接下来，我们试图回答这些问题。你会了解到：第一，利率既是短期货币市场的价格，又是资本市场的价格——它联结着两个市场；第二，从短期看，中央银行改变货币供给量可以影响利率，对真实经济产生作用；但在长期，改变货币供给量并不会改变均衡利率，因此不会对经济产生实际的促进，但可能产生坏的结果。

一、短期货币市场与资本市场

根据上一节第一小节的内容，我们可以在图 9.5 中绘制短期货币需求量与真实利率 r 之间的关系。

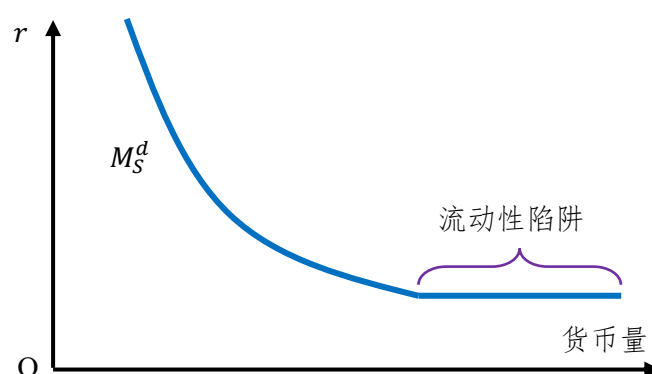


图 9.5：短期货币需求量与真实利率的关系

上图中比较有趣的一点是，需求曲线上出现了一段水平的区间，这被称为一段存在流动性陷阱（liquidity trap）的区间。那么，什么是流动性陷阱呢？事实上，这是凯恩斯对当年大危机时期货币政策无效的一种刻画。货币当局即使提高货币供应量，也无法继续降低利率，因为货币需求对利率的弹性已经处于无穷大的状态了。当经济处于流动性陷阱时，货币政策会失效，货币扩张也无法降低利率、刺激投资和增加总需求。因此，凯恩斯主张要改用积极的财政政策。

货币供给曲线在短期依然是一条垂直线，我们依然使用 M^s 来记。这样，货币短期需求和供给曲线绘制在一起，就会相交于一个短期均衡点（如图 9.6 左边所示）。有了短期货币市场模型，我们接下来就可以将其与可贷资金市场进行联系。

假如货币供给增加（ $M_0^s \rightarrow M_1^s$ ），那么在短期货币市场中，均衡利率就会下降（ $r_0 \rightarrow r_1$ ）。与此同时，可贷资金市场的投资量上升（ $L_0 \rightarrow L_1$ ），导致 GDP 上升，从而使得储蓄曲线右移（ $S_0 \rightarrow S_1$ ），并与投资曲线相交于货币市场决定的新的均衡利率 r_1 处。

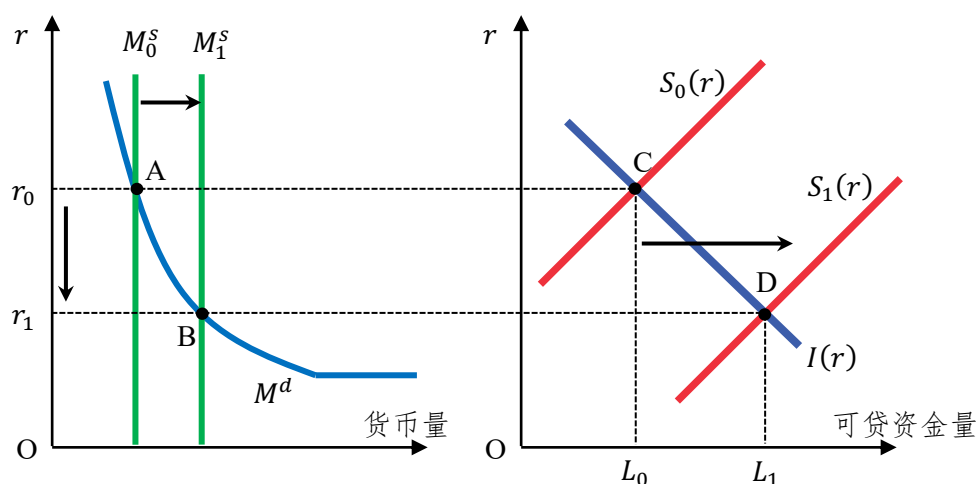


图 9.6：短期货币市场与可贷资金市场的联系

为什么最后可贷资金市场的均衡对应的利率恰好与货币市场决定的利率一致？这实际上来源于无套利条件。试想，若回购协议利率扣除风险溢价低于银行贷款利率，那么人们会跑到货币市场来融资，从而抬高了回购协议利率。我们简化了这个调整的过程，只看最初情况和最终（调整后）的结果。

图 9.6 中还有两个要注意的点：（1）私人投资的增加量（ $L_1 - L_0$ ）是真实量，并不一定等于货币供给量的增加量（名义量）；（2）一段水平的流动性陷阱区间在画图时可以不特别强调或画出来（大多数教科书选择将其隐去），因此下面的图像模型都不特意画出这一段区间。

二、长期均衡利率

上面讨论的都是短期效应，那长期呢？我们用图 9.7 来分析。

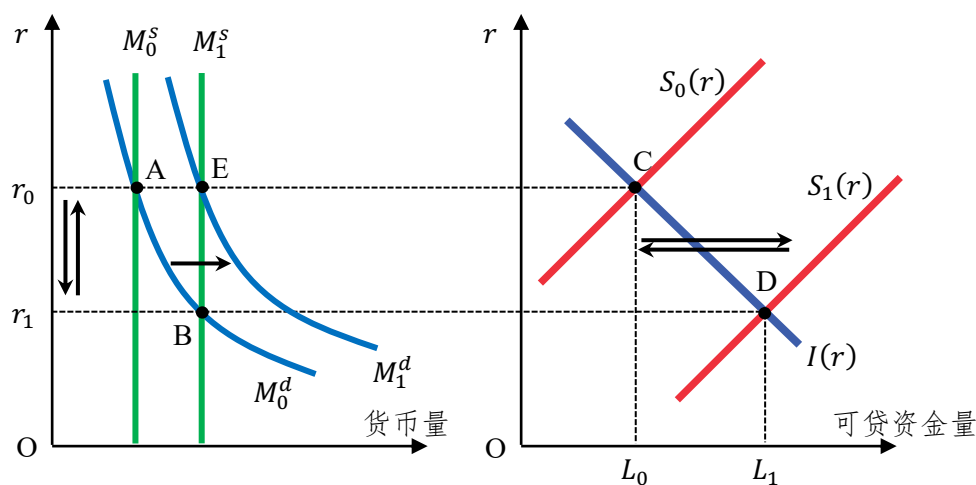


图 9.7：长期货币市场与可贷资金市场的联系

在长期，人们对货币的需求会受到收入 Y 和物价 P 变化的影响。真实收入高了，需要持有的货币自然也要增加；真实收入不变，但物价高了，也要增加手持货币。从图 9.7 中可以看到，当短期发生了一系列变化（ $M_0^s \rightarrow M_1^s$, $r_0 \rightarrow r_1$, $L_0 \rightarrow L_1$, $S_0 \rightarrow S_1$ ）之后，货币需求会上升（ $M_0^d \rightarrow M_1^d$ ），从而利率也会回升（ $r_1 \rightarrow r_0$ ）。与此同时，可贷资金市场的投资量回落（ $L_1 \rightarrow L_0$ ），导致 GDP 回落，从而使得储蓄曲线左移（ $S_1 \rightarrow S_0$ ），并与投资曲线重新相交于均衡利率 r_0 处。

从长期来看，货币增加并不会导致利率下降，也不会扩大投资，真实 GDP 也不会因为货币扩张而提升。那么什么变了呢？从图 9.7 左图可以看出，均衡从 A 点到 E 点，货币量提高了。这意味着，物价水平上升，或币值下降（图 9.8）。

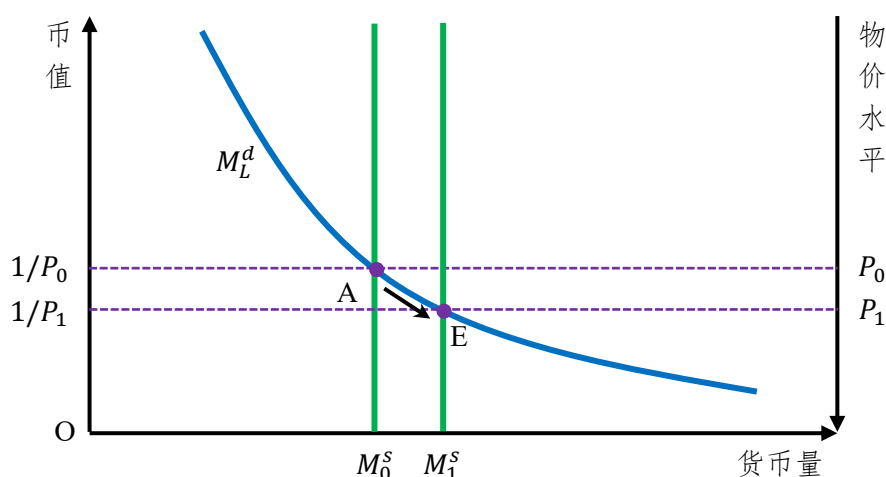


图 9.8: 货币扩张的长期影响

从图 9.8 中可以看出，由于真实利率在长期没有发生变化，因此我们在长期的图像模型中隐去了关于利率的讨论。⁶

拓展阅读 9.2

在 Krugman 和 Wells (2024) 的教材中，当讨论到图 9.6 和 9.7 中的模型时，他们将可贷资金供给曲线的移动归因于利率的变化。然而，这个说法的问题是， $S(r)$ 函数中的 r 是内生变量，不应该改变 $S(r)$ 的函数形式，即不应该导致 $S(r)$ 的位移。如果要移动 $S(r)$ ，应该也是因为真实货币供给的变化。在本讲义中，我们使用的说法是“与此同时”而不是“因此”。若非要归因，利率的变化应该是结果，而可贷资金供给的变化才是原因。真实货币供给在物价还没有上涨的时候是增加的，因此使可贷资金供给曲线右移。随后，物价上涨导致真实货币供给回到最初，因此使可贷资金供给曲线回到最初。

⁶ 事实上，根据公式 (9.5)， $M^d = \$Y \times L(i)$ 。在图 9.7 中，货币需求曲线的右移起初来自于 Y 的增加。随后，当 Y 回落时， P 随之上升，维持了短期货币需求曲线右移的位置。在物价上涨的过程中，名义利率高于实际利率；当物价上涨到稳定位置后，若通货膨胀率为 0，则名义利率与实际利率无异。在图 9.8 中，由于真实收入和利率被隐去，长期货币需求曲线起初因为真实收入的上升和利率的下降而提高（右移）了，因为 $L(i)$ 是 i 的减函数。随后，真实收入和利率回落，最后物价不再上涨时名义利率与真实利率无异，因此长期货币需求曲线又回到了原位。

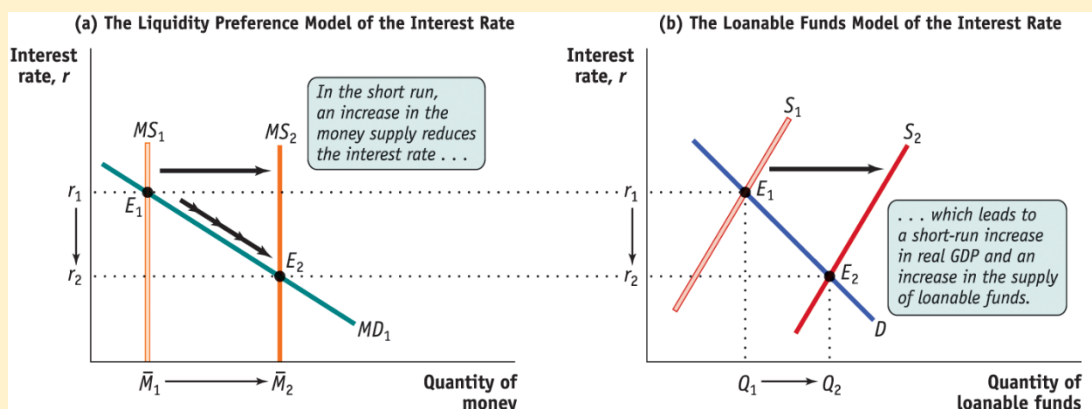


图 9.9：利率的短期决定

注：该图来自于 Paul Krugman 和 Robin Wells（2024）出版的宏观单行本《Macroeconomics》（第七版）第 15 章附录图 A-1。

Krugman 和 Wells（2024）另外一个问题是，当讨论到图 9.7 货币需求曲线移动时，他们只提到物价上涨，忽略了在调整过程中真实 GDP 也曾经有所增加。实际上，在调整过程中，货币需求增加同时来自 Y 和 P 的上涨，但最终只来源于 P 的上涨，因为 Y 先上涨后下降，如果假设货币中性， Y 最终不变。 $S(r)$ 位移回原位，来自于长期真实货币供给回归原位。

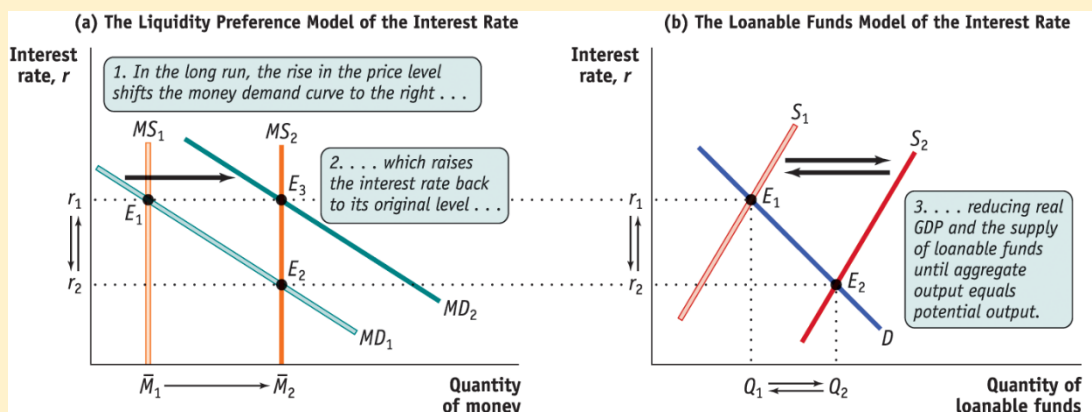


图 9.10：利率的长期决定

注：该图来自于 Paul Krugman 和 Robin Wells（2024）出版的宏观单行本《Macroeconomics》（第七版）第 15 章附录图 A-2。

第三节 货币市场均衡与货币数量论

本节内容将从货币市场的长期均衡谈起，并进而介绍经典的货币数量论。我们将探讨货币市场的长期均衡及其与货币数量论的关系。货币数量论为我们理解货币供应如何影响总体价格水平和宏观经济变量提供了一个重要的分析框架。基于这一经典理论，我们今后将能够更好地理解货币供应的变化是如何通过影响物价水平来

达到其最终效果的，并探索这些效应如何在长期市场均衡中显现。此后，我们还将讨论货币政策如何利用这些原理来实现稳定增长和控制通胀的目标，从而阐明货币数量论在现代经济政策中的应用与重要性。

一、货币市场的长期均衡与物价对货币供求的调节

我们在第一节第四小节中已经进行了初步讨论。在这一小节，我们再深入展开讨论其背后的机制。我们在图 9.4 的基础上加入一些新的标记，并绘制图 9.11。从下图可以看出，当长期货币需求因真实 GDP 增加而提高（ $M_A^d \rightarrow M_B^d$ ）时，需求曲线向右移动。在 P_A 的物价水平上，会形成货币的超额需求，真实量为 ACDE 的面积。这个时候，货币的供给真实量仅为 AEOG 的面积，而货币的需求真实量（CDOG 的面积）更大一些，市场出现了不均衡的情况。

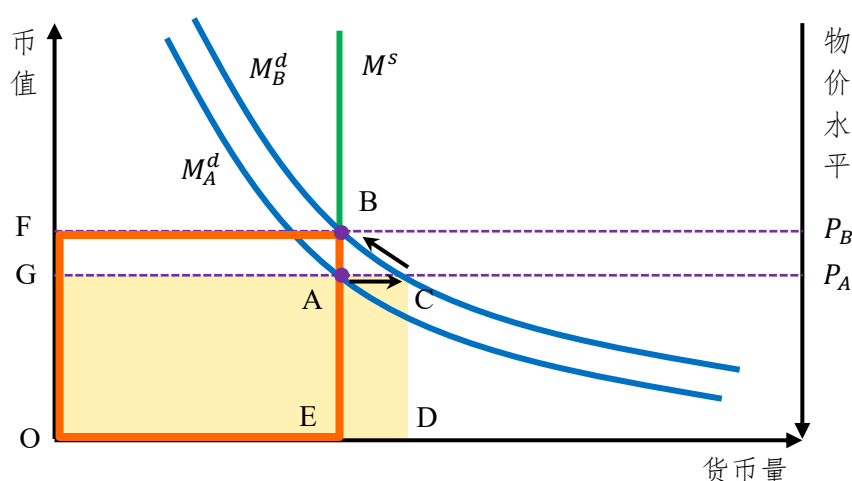


图 9.11：物价如何在货币需求上升后进行调整

这个货币的超额需求反映了新增真实产出带来的交易需求，而黄色矩形 CDOG 的面积则反映目前新的真实产出水平，超过了原先的真实产出水平（矩形 AEOG）。为了让人们能够使用同等的真实货币量完成交易，币值势必要有所提升，也即物价应该下降。当物价下降到 P_B 时，新的货币需求的真实量（BEOF 的面积）恰好等于新的货币供给的真实量，而且看起来也近似等于原先的货币需求的真实量（CDOG 的面积）。这样一来，市场再次回到均衡的状况——这个调节是通过物价实现的。

我们还可以在图 9.8 的基础上加入一些新的标记，并绘制图 9.12。从下图可以看出，当长期货币供给提高（ $M_0^s \rightarrow M_1^s$ ）时，会引致人们对产品与服务产生新需求（矩形 ABCD）。这在短期内确实也会有相应的生产活动来让人们得到满足，正如图 9.7 所讨论的那样。然而在长期，生产活动不足以支持矩形 BCOG 所对应的产品与服务需求，因为真实 GDP 会回落。最终，新的货币供给的真实量（矩形 BCOG）就超过了原先的货币需求的真实量（矩形 ADOG）。市场出现了不均衡的情况。为了让人们能够更多的真实货币量交易，币值势必要有所下降，也即物价应该上升。如果

物价上升到 P_1 ，则新的货币供给的真实量（ECOF 的面积）恰好等于新的货币需求的真实量（ADOG 的面积）。这样一来，市场再次回到均衡的状况——这个调节也是通过物价实现的。

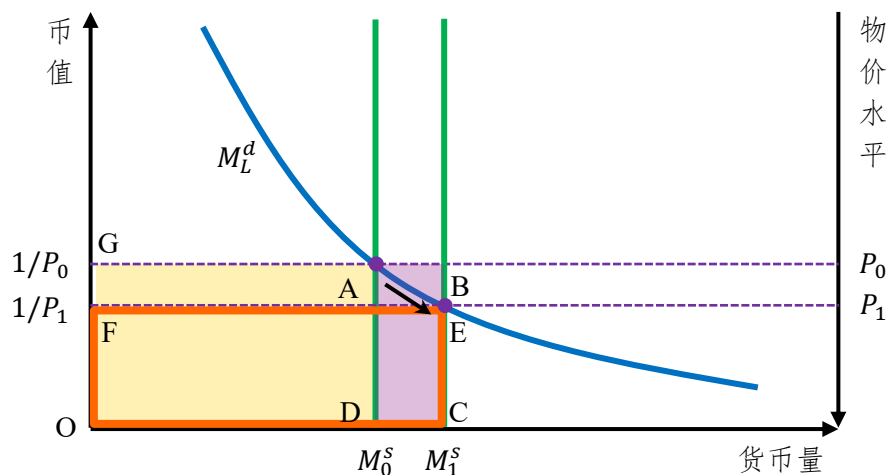


图 9.12：物价如何在货币扩张后进行调整

二、古典二分法

从上面的分析可以看出，货币数量的变化在长期只会影响物价水平和币值，而这正是货币数量论的核心。该理论最初被大卫·休谟（David Hume）甚至更早期的经济学家讨论，随后经过欧文·费雪（Irving Fisher）、阿瑟·赛斯尔·庇古（Arthur Cecil Pigou）、米尔顿·弗里德曼（Milton Friedman）等经济学家不断发展。如今，其相对合理的成分已经被现代经济学广泛接受。

我们先谈一谈休谟在 18 世纪中叶提出的古典二分法。休谟认为，所有经济变量应分为两类：

- **名义变量（nominal variables）**：按货币单位度量的变量。例如，一美分、一人民币元。
- **真实变量（real variables）**：按实物单位度量的变量。例如，一个鸡蛋、一支笔、一件上衣、一条电报。

例 9.3

在下面的变量中，哪些是名义变量，哪些是真实变量？

- 新住宅的价格（元/ m^2 ）
- 小王的年收入（元/年）
- 房价-收入比
- 猪肉的价格（元/斤）
- 一斤猪肉与一斤白菜的价格之比
- 失业人数（万人）

在上面的变量中，名义变量包括新住宅的价格、小王的年收入、猪肉的价格，而真实变量包括房价-收入比、一斤猪肉与一斤白菜的价格之比、失业人数。

房价-收入比之所以是真实变量，是因为它告诉我们一年（或一个月）的劳动可以换几平方米的新住宅。一斤猪肉与一斤白菜的价格之比则告诉我们一斤猪肉可以换多少斤白菜。这些都是真实变量。

事实上，货币也可以分为名义货币量和真实货币量。在上一小节中，我们已经使用了真实量的概念——它对应的是货币能够购买的产品与服务量，即 M/P ，其中 P 是用来衡量平均物价水平的物价指数， $1/P$ 是单位币值。那什么是名义货币量呢？实际上，我们一直说的货币量 M 就是名义量。比如，当你听到说我们国家增发货币四万亿元，那就是货币的名义量为四万亿。

休谟之所以要区分名义和实际变量，是因为在他看来，名义变量受到货币政策和货币制度变化的影响；但是，真实变量只在短期会受到这些影响，长期看则不会对货币进行响应。

拓展阅读 9.3

休谟的思想集中体现在如下两段话中。可以看出，他认为：在短期，增加货币可能有利于生产繁荣与经济增长；在长期，多发货币的结果只是引起物价上升。

“利息的低落一般起因于货币的充裕，但是……（货币数量）的增加所产生的效果只不过是提高了劳动和商品的价格……在这些变化的发展过程中，这种增加可以通过令人振奋的工业产生一些影响。但是，在价格稳定以后，……这种增加就毫无影响了。”

“尽管商品的高价是金银增加的必然结果，然而它并不是马上随着金银的增加而上升，而货币在整个国家流通之前，需要一些时间……据我看，只有在获得货币和价格上升之间的中间状况里，金银量的不断增加才有利于工业……”

——转引自 Friedman（1987）的《Quantity Theory of Money》，载于《新帕尔格雷夫经济学大辞典》第4卷，第4页，经济科学出版社1992年6月第一版

三、费雪交易方程式和剑桥现金余额公式

在传统的货币数量论中，有两个著名的公式，其中之一便是费雪交易方程式：

$$MV = PT \quad (9.7)$$

公式（9.7）来自于 Fisher（1911）的名著《The Purchasing Power of Money》（被翻译为《货币的购买力》）的第二版。在上式中， P 代表一个时期内的平均价格，而 T 则是这一时期内经过合适选择后的交易数量总量（因此也可以被写为 Y ）。因此， PT 就是这一时期内发生的交易的名义总额，相当于我们现在所说的名义 GDP。具体而言， PT 等于每种产品或服务的单价 p_i 与其交易量 q_i 的乘积之和：

$$PT = p_1q_1 + p_2q_2 + \cdots + p_nq_n = \sum_{i=1}^n p_iq_i$$

在公式（9.7）中， M 是货币量（存量）， V 是货币流通速度（velocity），是单位时间（如一年）内的周转速度。对于单笔交易来说， $V = 1$ ；然而，从一段时间交易总量的角度来看，一年内发生的交易总量 PT 让货币余额 M 周转了若干次，那么 V 就不等于 1 了。可以看出，该公式强调货币的交易媒介功能。

在图 9.11 和 9.12 中，我们实际上隐含假设了 V 是一个恒定的数。这样一来，当 $T = Y$ 增加时， $M/P = T/V$ 就会增加；若 $T = Y$ 保持不变，则 M/P 也保持不变。

剑桥学派的大师、马歇尔的继承人 Pigou 于 1917 年发表在经济学顶刊 QJE 上的文章⁷引发了第二个公式：

$$M = kPNy = kPy' \quad (9.8)$$

该公式其实是 Friedman 在《The New Palgrave Dictionary of Economics》（被翻译为《新帕尔格雷夫经济学大词典》）里的总结。在公式（9.8）中， y 是以不变价格计算的人均国民收入， N 是人口，于是 $y' = Ny$ 就是真实 GDP。 k 是货币量与收入的比值，例如：货币量是 10 万亿元，收入是 100 万亿元/年，则这一比值就是 0.1 年/单位货币。

需要注意的是， k 可以是一个“想要的”（desired）比值而不是实际的比值。在数学上， $k = \frac{1}{V}$ ，也即当人们只想把 $\frac{1}{10}$ 的年收入以货币的形式持有的时候，货币流通速度为 10 次/年，每单位货币流通 0.1 年。这个简单的数学变换产生了新的含义——该公式强调货币的价值贮藏手段功能，即有多少货币会在人们买卖之间成为购买力的临时栖身之所。

结合公式（9.7）和（9.8），我们可以写出

$$M = \frac{y'}{V}P \quad \text{或} \quad P = \frac{V}{y'}M$$

⁷ Pigou, A. C. (1917). The value of money. *Quarterly Journal of Economics*, 32(1), 38-65.

无论是费雪交易方程式还是剑桥现金余额公式，都指出 M 与 P 之间的正比关系。换言之， M 与 $1/P$ 之间存在反比关系，这也就是为什么前面长期货币需求曲线都画成了双曲线（hyperbola）的形式。当然， y' 和 V 的比例是可以变化的， V 也可以“包罗万象”，因此该关系是基本不会出错的。

货币数量论是一个恒等式。这是一个关于货币需求的恒等式，反映相关关系，但不反映因果关系。当 M 增加时， P 也增加，根据该公式，我们无法说谁是因谁是果——即不好说是谁在先、谁在后。如果我们进一步引入货币供给 M^s ，并使得 $M^d = M^s$ ，则可以进行因果推断。当我们提高 M^s 时，新的均衡必然由 P 的调整实现，因此 M 是因、 P 是果。

最后，我们简要谈一点剑桥现金余额公式与费雪交易方程式的区别。在费雪方程式中， V 一般被视为常数， $T = Y$ 也不取决于货币量，因而 M 的变化就只会引起 P 的同向变化。在剑桥现金余额公式中，实际上 k 是可以发生变化的：当 P 上升（即发生通货膨胀）时，人们口袋里会多放一点钱，因此 k 会上升（这样就意味着 V 下降）。如果我们不假定货币流通速度不变，则公式（9.8）实际上就变成了货币需求函数的一个表达式。

四、通货膨胀税与铸币税

最后，围绕着物价这件事情，我们讨论两种“税”。

（一）通货膨胀税

首先，通货膨胀可以被认为是一种税。为什么？征税无非是政府要求纳税人将自己的收入的一部分交出来，供政府来支配。为了征税，政府还得供养几百万人的稽税、征税队伍，成本很高。政府不如通过央行印钞票，然后发债借来这些钞票，并用这些新印的钞票给军队、官员、教师和医生发薪水开销，这就用不着麻烦税务局了，效果也是一样的。所以，通货膨胀、多印货币，相当于征税。

我们怎么度量通货膨胀税的规模呢？这里需要考虑两期之间大众手上的钱贬值了多少：

$$\text{真实通货膨胀税} = \frac{M_{t-1}}{P_{t-1}} - \frac{M_{t-1}}{P_t} \quad (9.9)$$

在公式（9.9）中， $\frac{M_{t-1}}{P_{t-1}}$ 是政府还没有大印钞票时的货币的真实价值，是 M_{t-1} 在 $t-1$ 期的真实购买力； $\frac{M_{t-1}}{P_t}$ 是这些钱在政府大印钞票后（ t 期）的今天的真实价值或购买力。由于印钞， $P_t > P_{t-1}$ ，因此必然有 $\frac{M_{t-1}}{P_{t-1}} - \frac{M_{t-1}}{P_t} > 0$ 。因此，通货膨胀税可以

理解为是由于物价上升使老百姓手中货币购买力减少的量就，相当于政府通过多发货币从百姓身上抽走的“税”。

（二）铸币税

我们再谈一谈**铸币税**（seignorage），这是政府用新发行的货币去购买产品、服务与非货币资产的收入。这与通货膨胀税有联系，但是也有区别。我们有两种衡量铸币税规模的方式：

$$\text{名义铸币税} = \Delta M_t = M_t - M_{t-1} \quad (9.10)$$

$$\text{真实铸币税} = \frac{\Delta M_t}{P_t} = \frac{M_t}{P_t} - \frac{M_{t-1}}{P_t} \quad (9.11)$$

公式（9.10）是说，政府新印多少钞票，这本身就是在征税；公式（9.11）则是说，这些新印的钞票能买多少产品和服务。可以看出，公式（9.11）和（9.9）视角不同。公式（9.9）是在说，百姓损失了多少价值；公式（9.11）是在说，政府得到了多少价值。

通过对比两个公式，不难发现政府得到的价值与百姓损失的价值并不相等：

$$\text{真实铸币税} = \underbrace{\frac{M_t}{P_t} - \frac{M_{t-1}}{P_{t-1}}}_{\text{货币总价值的变化}} + \underbrace{\frac{M_{t-1}}{P_{t-1}} - \frac{M_{t-1}}{P_t}}_{\text{真实通货膨胀税}} \quad (9.12)$$

通过对公式（9.11）的变换，我们发现真实货币税与通货膨胀税之间差了货币总价值（真实余额）的变化。根据费雪交易方程式，若货币流通速度和真实 GDP 没有发生变化的话，货币总价值的变化就是 0，那么真实铸币税与通货膨胀税就是等价的。然而现实中（根据剑桥现金余额公式），当政府印钞时，货币流通速度和真实 GDP 在短期会受到影响，而货币流通速度可能在长期也会发生改变。如果 V 下降了而真实 GDP 不变或上升， $\frac{M_t}{P_t} > \frac{M_{t-1}}{P_{t-1}}$ ，政府收来的真实铸币税还比百姓损失的多些，这就是通过印钞创造出来的额外的税收。

当然，公式（9.12）也警示政府不能印钞印得太过，否则会发生**恶性通货膨胀**。

当发生恶性通货膨胀时， $\frac{M_t}{P_t} < \frac{M_{t-1}}{P_{t-1}}$ ，真实铸币税要打一个折扣。极端情况下，人们对该国货币失去信心，即使政府持续印钞，人们也不持有了，那么甚至会出现 $\Delta M_t < 0$ （名义铸币税为负）的情况。此时，货币市场不处于均衡状态。

第四节 通货膨胀、反通胀与通货紧缩的成本

前面我们看到，货币供给多了，物价就会上涨，而在这个过程中有通货膨胀。接下来，我们要讨论通胀本身会带来的成本，以及为了消除通胀要付出的成本。

一、通货膨胀的成本

通货膨胀为什么会有成本？有的人可能会想，物价上涨了一倍，工资也涨一倍不就好了，只是所有数字都翻了一倍。根据古典二分法，真实变量在长期不会受到通货膨胀的影响。从费雪交易方程式出发，我们似乎也看不出通货膨胀有何成本，因为 $MV = PY$ ，如果 P 高一点， M 也相应提高一点不就好了？然而，这样的思考方式只看到了长期的结果，而忽略了短期的过程。比如，有人可能会对费雪交易方程式两边求全微分，则有 $\frac{\Delta M}{M} + \frac{\Delta V}{V} = \frac{\Delta P}{P} + \frac{\Delta Y}{Y}$ ，因此 $\frac{\Delta P}{P} = \frac{\Delta M}{M} + \frac{\Delta V}{V} - \frac{\Delta Y}{Y}$ 。若 $\frac{\Delta V}{V} = 0$ 和 $\frac{\Delta Y}{Y} = 0$ （货币中性），那么货币增发的速率就等于通货膨胀率，即 $\frac{\Delta P}{P} = \frac{\Delta M}{M}$ 。实际上，这个等式达成之前还发生了一系列的事情，而这些事情可能是有成本的。

总的来说，通货膨胀可能有七个方面的（并非互斥的）成本，下面我们对它们进行逐一介绍。

（一）皮鞋成本

通货膨胀的**皮鞋成本**（shoe-leather cost）是指由通货膨胀所引致的额外的交易成本。这个成本的名字是一种比喻，因为提到交易就会让人联想到人们东奔西走，导致皮鞋磨损。

与前面鲍莫尔-托宾模型相关的是，当通货膨胀发生时，人们会通过提高 N （多跑几趟银行）来减少现金持有（而增加其它保值或增值的资产的持有），由此就产生了额外的交易成本。

更进一步地，通胀促使人们频繁进行金融活动，如存取款、购买金融产品等，以逃避现金持有的贬值风险。这些活动不仅增加了个人和企业的财务管理成本，也可能对银行和金融系统造成额外负担。

另外，当通货膨胀不是均匀地发生在所有的产品和服务商时，人们可能还会花额外的精力去搜寻价格更合适的产品或服务。由于“货比三家”需要进行更频繁的物物比较，从而增加为了实现交易的成本。

1921-1923 年，德国发生恶性通货膨胀（hyperinflation），每个月价格上涨超过 50%，许多商人甚至会雇佣跑步健将，让他们在一天内多次帮他们把货币兑换成能保值的東西（如价值稳定的外币）。雇佣跑步健将来做这件事显然是一种成本。同时，由于银行交易次数激增，德国当时的银行不得不多雇佣三四倍的员工，这就导致实体经济中的工人不足，也是一种成本。

2018 年 4 月，根据委内瑞拉国民议会财政经济发展委员会发布的报告，与 2017 年同期相比，其通货膨胀年增长率达到 13779%。资料显示，委内瑞拉许多国民开始持有比特币、香烟而不是玻利瓦尔，甚至物物交易，这就增加了交易成本。

总而言之，皮鞋成本强调了通胀导致人们把资源耗在了不怎么创造价值的活动上面。

（二）菜单成本

简单来说，**菜单成本（menu cost）**是指因为改变标价而带来的真实成本。这个名称也是来自一个比喻，源自实体店铺需要重新打印价格菜单的成本，但实际上，菜单成本涵盖了所有因价格调整而产生的直接和间接费用。这些成本不仅包括物理上的重新标价成本，例如重新打印菜单或标签，还包括更广泛的行政成本，如重新设定销售系统、通知客户、重新培训员工关于新价格的信息等。

有的同学可能会问，现如今许多价格是在电子平台上标注，因此修改价格无非就只是更新一下数字，看起来也没什么成本？但是，电子系统的使用并不意味着成本的消失。例如，虽然在电子平台上更改价格无需支付打印费，但仍然需要支付员工的人工费用。即使现在有 AI 来代替人工完成很多任务，但开发或使用 AI 无疑也是有成本的，尤其是用于更新价格。

在高通胀环境中，菜单成本尤为显著，因为价格调整的频率增加，企业需要更频繁地更新和沟通价格变动。这不仅增加了运营成本，还可能影响企业的定价策略和利润率。企业在决定是否调整价格时，也需要权衡潜在的收益与菜单成本之间的关系。如果预期通过价格调整能显著增加收入或市场份额，企业可能就会决定承担这些成本。否则，它们可能会推迟价格调整，以减少开支。

菜单成本是价格粘性的一个重要原因。价格粘性意味着价格不会立即响应市场条件的变化——这是因为企业为了避免频繁变动价格所带来的成本，可能会选择在一段时间内保持价格不变。而这何尝没有成本？

当恶性通货膨胀来临时，许多人力物力不得不投入到时刻更新价格这项任务，有时人们干脆放弃使用本国货币标注价格，而使用他国货币标注价格又会带来额外的计算成本。1980 年代，以色列房地产业开始采用美元来计价，尽管交易结算还是使用谢克尔（shekels）。2009 年，在恶性通货膨胀发生后，津巴布韦甚至允许人们使用外币交易。

（三）价格尺度成本

通货膨胀的**价格尺度成本**或称**计价单位成本（unit-of-account cost）**指通货膨胀使得货币作为计价单位的可靠性受损所带来的成本。

现如今，契约合同中的交易是用货币计价的，经济决策也是基于用货币衡量的变量来制定的。我们在做经济分析时，通常需要基于货币计量的变量来判断企业的表现好坏，甚至预测企业未来的发展等等。通货膨胀的发生显然让货币变成一个不那么“靠谱”的价格尺度。许多经济学家认为，这个损失主要体现在经济决策质量的下降（这是“坏数据”导致的）。就好比拿一把错误的尺子参加数学考试，就可能答错一道最基础的几何题。

若这把尺子错得很有规律，比如它的刻度与准确尺子的刻度呈现 1:2 的关系，那么也未必会有损失。难点在于，这把尺子与准确尺子的关系是未知的，因为通货膨胀率具有一定的不可预测性。这样一来，决策者“算错”的概率就增大了。

价格尺度成本对于税收体系来说尤为重要，因为通货膨胀会让收入的衡量遭受扭曲。例如，当美国的通货膨胀率为 10% 时，一个用美元计价为 10 万的资产一年后以 11 万的标价卖出，这一万美元的差价将成为“虚幻的利润”（phantom profit）。如果税收体系对这个虚幻的资本利得征税，则可能劝退投资者。1970 年代，美国也经历了高通胀的年代，一些商人就不愿意投资在生产率高的项目上，因为他们认为自己要为“虚幻的利润”纳税。

总之，价格尺度成本关注的是货币作为计价单位的稳定性和可预测性问题，而菜单成本则是指改变价格标签所产生的直接和管理成本。

（四）资源错配

资源错配（resource misallocation）是价值尺度成本的一种具体表现形式。价值尺度信度的下降还来源于不同产品的价格调整频率存在差异，从而导致用货币度量的相对价格发生的变化不来自于供求的相对变化。这种相对价格的变化使各个市场出现供求不等，从而偏离了资源最有效率的配置。

我的一项近期研究发现，中国过去几十年来房地产市场价格远超其他行业增长就引起了企业资源的错配。这里，房价上涨有一定程度是因为通货膨胀，并且中国房地产相比其它行业可能是对通货膨胀响应较强烈的一个行业。

拓展阅读 9.4

房价上涨可以通过多种方式导致资源错配，影响经济的整体效率和生产性。资源错配意味着资本、劳动力和其他资源没有被用在产出最大化或最有效率的方式上。以下是房价上涨可能导致资源错配的几种机制（欢迎对这个话题感兴趣的读者补充）⁸。

第一，投资偏向的扭曲。当房价持续上涨时，个人和企业可能将更多的资金投资于房地产市场，寻求资本增值，而非投资于其他可能带来更高生产性或创新的行业。这种资金流向的改变可能导致对制造业、科技、教育和健康等关键领域的投资不足。房地产市场的吸引力可能导致资本从生产性业务转向投机性房地产交易，减少了对创新和技术进步的资金支持。

第二，劳动力流动性降低。房价上涨提高了居住成本，尤其是在城市和经济中心。这可能限制了劳动力从经济低效地区向更高效地区的流动，因为工人可能无法负担迁移到这些地区的成本。结果，一些地区可能面临劳动力短缺，而其他地区则可能有过剩的劳动力。高房价可能使得某些地区的企业难以吸引和保留员工，从而影响这些地区的经济增长和发展。

⁸ 庄晨近期发表的论文对这一问题有补充讨论，感兴趣的同学可以阅读：Yu, F., Li, N., Zhuang, C. C., & Chen, J. (2024). Can rising urban house prices actually limit the outward FDI by firms in a home country? A story from China. *Journal of International Money and Finance*, 147, 103164.

第三，消费支出不足。房价上涨导致住房成本占家庭收入的比例上升，减少了消费者在其他商品和服务上的支出，这就可能导致整体需求下降，影响广泛行业的收入和增长。家庭为了支付高额房贷可能不得不削减教育、健康和休闲等方面的支出，这些都是影响长期经济增长和生活质量的关键因素。

第四，政府资源分配问题。政府可能需要增加在住房补贴和社会福利上的支出来缓解房价上涨带来的影响，这可能会减少对教育、基础设施和其他公共服务的投资。

房价上涨如果不受控制，不仅可能造成经济资源的错配，还可能加剧社会不平等，限制经济的潜在增长。因此，理解房价上涨如何导致资源错配，并采取适当的政策措施来应对，对于维持经济的健康和可持续增长至关重要。

（五）税收扭曲

这是价值尺度成本的另一种表现形式。上面已经提到了“虚幻的利润”，这里就具体给出一个例子。

假设这个世界上有两个经济体，分别是 A 和 B。两个经济体的政府都征收利息税，且税率都为 25%。假如真实利率为 4%，经济体 A 没有通货膨胀，但是经济体 B 有 10%的通货膨胀率，那么，两个经济体的税后真实利率分别是多少呢？

表 9.1：一个税收扭曲的例子		
	经济体 A	经济体 B
通货膨胀率	0%	10%
税率	25%	25%
税前名义利率	4%	14%
税前真实利率	4%	4%
税后名义利率	3%	10.5%
税后真实利率	3%	0.5%

从表 9.1 中可以看出，在一个通货膨胀率较高的经济体，即使税率与其它经济体相同，也会导致不同的税后真实利率，这是由于对“虚幻的利润”征税导致的。较低的真实利率会减少储蓄，即减少可贷资金的供给，从而不利于长期经济增长。

一个方案是将税法指数化（indexation），即根据物价水平（如 CPI）动态调整税法——事实上，该做法已经在个人所得税中开展，但资本利得税等其它税种还未得以普及。另外，指数化让原本复杂的税法更加复杂。

（六）混乱与不便

这也是价值尺度成本的一种表现形式。财务会计在核算公司的收益时，若存在通货膨胀，投入与产出可能不可直接比较，尤其是生产周期比较长的公司的收益。这种不便比较的情况可能导致统计混乱，进而影响金融市场对资源的配置。

（七）随机财富再分配效应

前面讨论的第四、五和六点成本在通货膨胀稳定、可预期的经济体中都是存在的。有一种成本可能只由意外的、不可预期的通货膨胀造成：债权人（creditors）和债务人（debtors）无法合理预期通货膨胀时，有一方会得益，而另一方则受损，而这就是财富从一方向另一方的转移（再分配）。这实际上也可以算作价值尺度成本的具体表现形式之一。

从各国的数据中可以发现一个规律：平均通货膨胀率较低的经济体的拥有比较稳定、可预测的通货膨胀率；而通货膨胀率较高的经济体的通货膨胀率则波动大、不好预测。我们几乎看不到哪一个国家的通货膨胀率稳定地维持在一个高的水平。这个现象指出了通货膨胀的另外一个成本：当一个国家希望推行高通胀的货币政策时，不仅要承受上面提到的可预期的通胀成本，还要承受不可预期的通胀成本。而这个不可预期的通胀成本，来自于财富再分配的随机性。它导致了部分人损失，并可能带来经济效率的损失。

二、反通货膨胀的收益与成本

通货膨胀，尤其是高通货膨胀，对经济不利，因此，政府通常会为反通货膨胀（disinflation）付出努力。然而，反通货膨胀本身既有收益，也有成本。

（一）反通货膨胀的收益

马丁·费尔德斯坦（Martin Feldstein, 1939-2019）于 1997 年⁹进行了一项研究，发现在当时税法和通胀的相互作用下，即使通胀率在低水平（如 2%）的情况下也会造成不小的福利损失。若将通胀率从 2%降低到 0%（即“价格稳定”水平），能够持续获得等价于每年 GDP 的 1%的收益。

反通货膨胀通过避免或减少上一小节提到的通货膨胀的成本（无谓损失）提高了社会福利。除此之外，还有什么收益呢？

经验表明，当高通胀成为一种“常态”以后，反通胀将变得非常困难。因此，积极反通胀、将通胀控制在低水平能够避免未来的“麻烦”。这或许也可以算一种收益。

（二）反通货膨胀的成本

Feldstein（1997）也计算了将通胀率从 2%降低到 0%的“一次性”成本，相当于 5%的（当期）GDP。事实上，若要将通胀率从更高水平（如 10%）下降到低水平（如 2%）要付出的成本或许更高。

⁹ Feldstein, M. S. (1997). The costs and benefits of going from low inflation to price stability. In *Reducing Inflation: Motivation and Strategy* (pp. 123-166). University of Chicago Press.

我们在之后的章节会具体展开讨论反通胀的成本，这里只是简要先提一提。最主要的成本就是让经济降温，使得失业率上升。

三、通货紧缩的成本

通货膨胀是坏事，通货紧缩（deflation）或许更糟糕。通货紧缩就是物价下降，通货膨胀率为负数。中国在 1963-1967、1998-2002 年间就发生过这样的现象。物价下降听起来是件好事，因为东西变便宜了。然而，从宏观经济学的角度来看，这也并非没有成本。为什么呢？

首先，前面讨论的通货膨胀的成本也适用于通货紧缩，只是方向反过来（只要价格改变，就会有皮鞋成本、菜单成本和价值尺度成本等），比如通货紧缩也可能导致随机财富再分配。

其次，通货紧缩可能反映更深层次的宏观经济问题，失业率进一步上升，生产活动进一步减少。由于名义利率最低往往是 0，通货紧缩可能意味着真实利率较高，那么就很少人愿意消费、投资实业了。

这里重点提一提两种效应。

第一，债务紧缩效应（debt deflation effect），也即债务人在物价下降时会缩减开支而对经济产生影响。为什么呢？因为债务人要按时还一笔事先约定好的欠款，而这笔固定的金额在通货紧缩下会具有越来越高的真实价值；债务人能否还得起债取决于他能否有足够的“生意”获得足够的收入，但在通货紧缩下“生意”越来越难做，收到的钱的名义量越来越少。债权人为了还得起债，只能缩减开支。债权人越是紧缩开支，经济就越是下滑，社会总需求下降，物价进一步下跌，然后债权人又继续紧缩开支……这就进入到了一个恶性循环中。

第二，预期紧缩效应，指人们预期通货紧缩将会持续下去而改变行为的影响。当消费者预期未来价格会继续下降时，他们可能会推迟购买决策，以期将来以更低的价格购买相同的商品或服务。这种预期减少了即时消费需求，推迟了购买决策。同样，企业在面对预期的未来成本下降时，可能会推迟投资决策，如购买新设备或扩建设施，希望在成本更低时进行这些支出。而持续的价格下降可能会影响到整体经济信心，人们对经济前景持悲观态度，担心失业和收入下降，从而减少了支出和投资。

当然，除了上面的两种效应以外，还可能其它效应。例如，可能会导致银行信贷紧缩。在价格持续下降的环境中，银行可能面临资产价值下降和坏账风险增加的问题。因此，银行可能会减少贷款发放，增加借款难度，从而进一步限制了企业和消费者的支出和投资能力。

日本自 1990 年代初以来一直面临长期的通货紧缩和经济停滞问题，这导致了日本央行采取了一系列非传统的货币政策措施，包括实施负利率政策。日本的负利率政策和持续的通货紧缩状态带来了多重经济挑战，尽管有其短期内的正面影响，但长期而言，这些政策可能导致更多的经济和金融问题。我们将在之后的章节再讨论货币政策的相关问题。

拓展阅读 9.5

弗里德曼规则（Friedman rule）是以 1976 年诺贝尔经济学奖得主米尔顿·弗里德曼的名字命名的，旨在通过操控名义利率来最大化社会福利。他认为最优的货币政策应该设定为名义利率为零甚至略微负值，以最小化持有货币的机会成本——即放弃的利息。通过将名义利率降至零，货币和债券（或其它名义资产）变成了完美的替代品，在理论上消除了持有货币这种非生息资产的低效率。

实施弗里德曼规则意味着对通货膨胀采取独特的方法。为了实现名义利率为零，真实利率（经通货膨胀调整后的利率）必须通过负通货膨胀率来抵消，也即需要有轻微的通货紧缩。将利率降低至零可以增加流动性，使得货币在没有放弃利息的情况下随时可用于交易。随着持有成本的减少，消费者和企业可能更倾向于进行支出和投资，可能促进经济活动。当然，该规则可能固有地导致通货紧缩，这在经济上有其自身的一系列影响，包括增加真实债务负担和可能的经济增长缓慢，如果管理不当的话。

虽然在理论上具有吸引力，弗里德曼规则也面临批评和怀疑，特别是关于其实际影响的批评。持续的通货紧缩可能导致消费和投资延迟，因为市场预期价格将进一步下降，这可能抑制经济增长。另外，接近或达到零利率时，货币政策的有效性可能受限。

课后思考题

请阅读下面这段材料，回答相关问题。

官方上来说，电子资金转账公司 PayPal（也是流行的手机支付服务 Venmo 的所有者）并不被视为银行。相反，监管机构将其视为货币传输机构，这是一种将你的资金发送到某个地方而非持有并保管它的实体。

然而，随着用户在其 Venmo 和 PayPal 账户中积累了大量资金，这种区分开始显得有些值得质疑。特别是 Venmo 用户，他们似乎常常愿意让收到的付款留在账户中，直到资金被支出，就像人们曾经将现金保存在钱包中一样。因此，2022 年 PayPal 的账户总额估计超过 365 亿美元。如果这些资金被视为银行存款，PayPal 将被视为美国最大的 60 家银行之一。

乍一看，将大笔资金留在 PayPal 账户中似乎违反直觉，原因有二。首先，这些账户没有联邦存款保险的保护。其次，它们不支付利息。但仔细考虑后，这种行为在经济上是合理的。人们通常只会在其 PayPal 账户中保留他们财富的一小部分，因此缺乏联邦存款保险是可以接受的风险。而且，在撰写这段材料时（2022 年 3 月），银行账户的利率非常低（约为 0.06%），因此损失这部分利息以避免在银行账户和 PayPal 或 Venmo 账户之间反复转移资金的麻烦是合理的代价。

结果是，许多人的行为就像《华尔街日报》引述的一位用户那样，她现在会等一段时间后才将资金从她的 Venmo 账户转出到她的常规银行账户：“我开始有意识地将我的钱在那里多留一会儿。”

但是，PayPal/Venmo 或类似的东西会开始对传统银行业造成重大影响吗？一些分析师这么认为。然而，其他人则暗示，像 Zelle 这样由传统银行开发的支付系统，以及最终上升的利率，将吸引客户回到传统的银行存款。时间会给出答案。

- 1、PayPal 账户没有被计入货币供应量中。它们应该被计入吗？为什么或为什么不？
- 2、2010 年，美国只有大约 25% 的手机是智能手机。到了 2023 年，几乎每个人都拥有智能手机。这种情况如何与 PayPal 的故事相关联，又是如何符合更广泛的货币历史模式的？
- 3、解释美国 FOMC 提高利率的行动将如何影响 PayPal 及类似服务的未来？

作业题

- 1、下表给出了 7 个国家 2010 年的通货膨胀率和 2011-2021 年的平均通货膨胀率的数据（来自 IMF）。请根据数据回答问题。

国家	2010 年通货膨胀率	2011-2021 年平均 通货膨胀率
巴西	5.04%	5.88%
中国	3.18	2.37
法国	1.53	1.07
印度尼西亚	5.13	4.17
日本	-0.73	0.47
土耳其	8.57	11.06
美国	1.64	2.00

- a. 根据平均通胀率和菜单成本之间预期的关系，请使用 2011-2021 年的平均通胀率，按菜单成本降序排列各国。
- b. 按照对 2010 年十年期贷款的借款人最为有利的通货膨胀率顺序排列各国。假设这些贷款是在预期 2011-2021 年的通货膨胀率将与 2010 年的通货膨胀率相同的情况下达成的。
- c. 相比在日本的贷款人，那些在日本借入十年期贷款的借款人总体上是获益还是亏损？请解释。