

# 社会主义和资本主义

## 关于短缺和过剩的市场理论研究

雅诺什·科尔奈

### 1. 导论<sup>①</sup>

#### 直观印象

打开英文版维基百科有关“短缺经济”（shortage economy）的网页，你会看到一张 1980 年代的照片，人们在波兰的一家食品商店门外排起长队。<sup>②</sup>而今天在波兰购物的人看到的则是琳琅满目的货架、来自国内外的品种繁多的商品。

\* 本文自英文版翻译而来。原作为匈牙利文（Hiánygazdaság – Többletgazdaság），收录于作者的研究文集《对资本主义的思考》（Gondolatok a kapitalizmusról），由布达佩斯的 Akadémiai Kiadó 出版。

<sup>①</sup> 本文是对作者以前的一篇论文的修改和完善：Közgazdasági Szemle (Kornai, 2010b)。感谢几位同行的帮助，特别是 Attila Chikán 和 Zsolt Mátysz 在数据收集和问题整理方面的工作。还要感谢诸多同行的宝贵意见：Daniel Brooks, Zsuzsa Dániel, János Gács, Dóra Györfly, Mária Lackó, Aladár Madarász, Zsolt Mátysz, Ágnes Nagy, Éva Palócz, András Prékopa, András Simonovits, Domokos Szász 和 István János Tóth。特别感谢英文译者 Brian McLean 一贯的出色工作。我高度评价几位亲密同事在信息、数据和文献搜集以及文章编辑上给予的细致合作，他们是 Rita Fancsovits, Ildikó Magyar, Andrea Reményi 和 Katalin Szécsi。感谢布达佩斯学院提供的思考环境和有效支持，让我克服了写作中的若干困难。

<sup>②</sup> 参见 Wikipedia (2010)。

1999 年我到中国访问，以迎接我的《短缺经济学》第二版的出版。该书的编辑和几位熟悉其内容的学生带我到百货商场和食品店参观，其中一位提醒说，“现在您应该写一本新书，讨论过剩经济，而不再是短缺经济”。

我不会真的去写这么一本书，但我希望在一篇论文的框架内，把有关过剩经济理论的基本思想勾勒出来。其核心思想可以表述如下：正如短缺经济是社会主义经济制度的典型特征那样，过剩经济（surplus economy）恰恰是资本主义经济制度的标志。<sup>①</sup>

按照熊彼特的观点，研究工作始于“分析前的认知过程”（pre-analytical cognitive act），即为分析工作准备原始材料。他称其为洞察力，洞察力可以带来对事物的新认识（Schumpeter, 1954, 第 41 ~ 42 页）。我相信，这种分析前的洞察力包含在好奇心中，研究社会主义短缺经济经济学家就是带着这样的好奇心来仰视资本主义市场上的丰富供给。我认为，直观印象非常重要。或许正是通过把资本主义与另外一种经济制度（在某种意义上是其对立面）进行对比，让我获得了某些特别的认识，而那些习惯于生活在资本主义中的经济学家同行们则没有注意到，也未能将他们从传统观点中解放出来。

和我讨论这一议题的主流经济学家通常只赞同我的部分看法。他们很容易相信社会主义制度是一种短缺经济，但是，资本主义制度又为什么应该被称为“过剩经济”呢？商店和工厂仓库里肯定有额外的存货，产能也往往有富余，但这些过剩是必需的。生产商会生产那么多产品，贸易商会囤积那样多存货，是因为市场均衡的需要。社会主义经济并不处于均衡状态，而资本主义经济是均衡的。虽然存在上下波动（目前的波动幅度就很大），但均衡毕竟是长期趋势。

我们之间仅仅是词汇上的口舌之争吗？主流经济学家们所说的市场均衡，是否等于我说的过剩经济？两者是否没有本质区别？那样的话，我应该就此停笔，因为名称本身并不值得争论。

虽然他们看到的和我看到的是同一现象，但双方理解的含义却不相同，这是因为观点的差异。针对资本主义经济制度的现实表现——大量闲置的产能、丰富充实的货架和四处寻觅工作的劳动力——我给出的是全然不同的解释。每

<sup>①</sup> 我很犹豫是否采用这一表达方式。在英语中，除了“surplus”一词外，我还考虑过其他词汇，如 plenty、glut 和 slack 等。“surplus”似乎最适合这里描述的现象。

个生产商或贸易商都在独立进行决策，根据自身的利益需要，确定具体的存货和产能储备水平，但如果把商店里所有待出售商品和产能储备可能生产的所有产品加起来，将大大超出全社会可能购买的全部商品的数量。宏观经济学家可能认为，总就业量是已接近自然就业率的水平，可很多人依然感觉找不到工作，如果他们能够参与就业，社会的产出水平也会得到改善。

如果我的看法正确，资本主义制度的确是过剩经济，那将具有重大意义。它会影响到资本主义制度下的经济当事人的行为模式，会给资本主义的许多特征带来新的认识。从这个角度讲，本文的分析可能有助于更全面地从实证的角度去理解资本主义。<sup>①</sup>

此外，我还会得出一些规范角度的结论。主流经济学家们对于资本主义经济中的产能过剩、存货增长或过度供给感到不安，将这些资源视为浪费。但我认为，过剩经济虽然也有几方面的有害副作用，却是资本主义制度最大的优越性之一。

### 为澄清概念所做的初步尝试

人们对科学研究的要求是，采用具有明确定义的概念，推导出一系列论点。绝大多数经济学家对“供给”和“需求”这样的基本概念没有什么异议，萨缪尔森著名的经济学教材的早期版本曾在其中一章引用过如下箴言：“你甚至可以把鹦鹉训练成政治经济学家，它只要学会两个词汇就行：供给和需求”[Samuelson, 1980 (1948), 第52页]。然而这两个概念却给我带来了一些困惑，微观经济学中的其他几个标准概念也同样如此。只有在对我探讨的现象进行大致描述之后，才能对这些概念进行批判性的分析，得出完整的概念体系。我将分两个阶段来完成上述工作，在引言中对有关概念的解释并不十分严格，我知道大多数读者会按照通常的含义来理解这些概念，与我打算介绍的内容有差别。因此，这里首先对下列概念做些解释。

• 我所说的“过剩”大致等于标准的经济学理论所说的“过度供给”，也就是代表供给大于需求的情况。

<sup>①</sup> 这里使用的“实证”(positive)一词是科学哲学上的概念，强调对现实的客观描述和解释，与“规范”(normative)研究相反。

• 我所说的“过剩经济”，如果必须用标准化的术语来描述，我希望称其为“过度供给经济”。与之对应的“短缺经济”，则是“过度需求经济”。前者是一种普遍存在过度供给现象的市场状态，后者是一种普遍存在过度需求现象的市场状态。

在做完上述这些解释之后，我希望已与经济学家读者们建立起了正确的沟通，让大家能够跟上我的论述思路。在下文继续解释自己的理论时，我还将对有关概念以及它们涉及的测算问题给出更精确的定义。

### 本文的研究方法在经济学研究中的位置

本文的每句话恐怕都有前人的成果为基础。经济学的各个学派都有很多言论涉及市场问题，自我感觉最好的是主流经济学家：他们毫不怀疑现有的概念框架和局部现象分析背后的理论基础，在教科书和研究文献中将其视为既定的前提。

本文的研究可以视为挑战：在某些重要问题上，我提出了与主流学派不同的观点和思考方法。但遗憾的是，我不能简单地加入某个现成的、与主流学派有根本或部分差别的“异端”学派，在他们现有的理论架构上填上一笔，也不能直接沿用他们的术语。

不过我并不感到孤独。尽管我的整体思考框架是源自独创，但其中很多重要元素却来自前人的研究。在某些问题上，我的主张与主流经济学中的法则一致或相似，在其他一些要点上，则与这个或那个非主流学派存在联系。我会很高兴地在恰当的地方指出这些借鉴和共识之处。

本文的写作有几方面的诱因。首先，目前仍未平息的经济危机促使所有负责任的经济学家反躬自省，我们必须搞清楚用于分析经济运动的工具是否合适。对我本人而言还有其他几个动机。50多年前，在准备自己的博士论文《经济管理中的过分中央集权化》(Overcentralization in Economic Administration, Kornai, 1957)时，我开始认识到，长期存在的短缺是社会主义经济制度面临的一个根本问题。在我后来的著述中，有关短缺、过剩、均衡和非均衡等问题也频繁出现例如<sup>①</sup>。这方面的议题看来将伴随我一生的思考。不过，本文的论

<sup>①</sup> Anti-Equilibrium, Economics of Shortage, The Socialist System (Kornai, 1971, 1980, 1992) .

述并不仅仅是重提我过去的思想。这个议题仍在继续，而我的理论知识和实践经验在此期间却有了成长。我希望自己的分析工具在不断更新和打磨之后已有所进步，我希望看到，这篇文章重塑了我早期的分析框架，其构造更加精细，某些观点得以修正，许多细节被彻底地重新加工。

我早年的研究主要是关注社会主义经济制度的运转，将资本主义经济制度主要作为其对立面进行比较研究。如今的研究重心有了转移，主要是分析资本主义。本文不但试图开发新的分析工具，而且在分析对象上超越了以前的作品。<sup>①</sup> 我将试图把半圆形的作品完善，把缺失的部分用曲线补充上。<sup>②</sup> 我希望说服自己（或许也包括读者），本文所使用的方法和工具也能够适用于资本主义市场经济。

## 研究对象及其结构

这里讨论的主要是资本主义经济制度，仅仅是出于对比的需要，才会涉及社会主义制度的某些内容。<sup>③</sup>

现代资本主义通常是一种混合经济：既有私人所有制和市场机制调节的领域，也包含相当数量的公有制和官僚机构调节的活动。

本文的前五章主要讨论市场领域的问题，第6章讨论非市场领域的问题。

研究仅限于产品和服务市场以及劳动力市场，也就是实体经济部分，没有详细论述货币经济的部分，即金融业、货币交易、信贷和金融资产投资等，也不涉及政府的财政和货币政策等。本文对这方面的问题仅会点到为止。

我承认，在讨论产品、服务和劳动力市场时，完全忽略金融产业（如货币、信贷、利率、财政和货币政策等）及其与实体经济之间的相互作用几乎是不可能的。但我依然决定对二者予以明确区分。本文的写作篇幅有限，并不是要撰写一本有关整个资本主义制度的巨著。

<sup>①</sup> 本文提到了我以前的很多著述及其内容，读者们却未必完全清楚。因此我会试图进行简要介绍，以便读者能够跟进。我还必须指出，《社会主义体制》（*The Socialist System*, Kornai, 1993）一书将有助于对本文的理解，目前的研究可以视为其后续。

<sup>②</sup> 不幸的是，某些关键的部分仍缺失。下一节将会强调，本文未涉及货币领域的问题，对市场运转的解释是不完整的。

<sup>③</sup> 我在本文要表述的有关社会主义的内容，都包含在《社会主义体制》一书中。

虽然把资本主义的一些关键部分排除在了论述范围之外，但需要涉及的领域仍然非常广泛。对一些重要而复杂的问题，我的论述也必须相当简略。这些问题都有大量的研究成果，每个都值得进行专题探索。

为方便读者，这里介绍一下本文的内容结构。

第2~6章讨论基本场景，其中，第2~5章描述现象及其产生和重复产生的因果关系，第6章讨论其影响并对其进行评价。第7章讨论基本场景之外的特殊场景。文章最后，读者们自会清楚我所说的基本场景和特殊场景的含义。

## 2. 产品和服务市场：过剩的再生产机制

### 经济史上的一个案例：美国的电信系统

我首先介绍经济史上的一个案例，它所代表的现象将在下文进行普遍意义上的讨论。这是关于美国电话服务的发展历程（Grover 和 Lebeau, 1996; Atkin、Lau 和 Lin, 2006），故事从1980年代初开始，当时几乎所有通话都在固定电话上进行，庞大的美国电话电报公司（AT&T）处于接近完全垄断的地位。

此时首先在技术上产生了重大革新，移动电话在1980年代后期快速普及（King 和 West, 2002），<sup>①</sup> 后来还出现了更多技术革新，这些将在下文再讨论。

1984年，市场结构也发生变化，根据美国的联邦反垄断决议，AT&T公司被逐步拆解。此后，随着公司的合并与分拆，市场的进入和退出，市场参与者的身份几经改变，但市场构成的基本布局却保留下来。这个市场既不是垄断市场，也非完全竞争市场，而成为少数几家实力强大的竞争对手形成的垄断竞争市场。

市场结构的这个特点与产业发展互为因果。创新接踵而至，移动电话不但能传递声音，还能发送文字信息和图片。便于携带的手机上增添了各种新功能：文字信息、图片和视频的传送，图像的存储以及计算器等等。

另一个创新方向则与传统的电话设备无关，通过 Skype 软件和其他类似网

<sup>①</sup> 与美国领先、欧洲追赶的通常规律不同，移动电话的普及表现出相反的方向。

络，人们可以用计算机进行语音甚至视频通话。

计算机和电话的功能已走向融合，还表现在其他各种形式上，如流行的 iPhone 手机和 iPad 便携电脑及其他竞争厂商的产品，都可以实现手机通话和其他信息传递与娱乐功能。每次创新热潮都给这个产业带来了新的生产商和服务提供商，有的老牌企业则被边缘化甚至淘汰出局。不过，市场的主要特征依然没有改变：少数大企业掌握统治权，彼此竞争、补充与合作，也极力打击对手。

从供给方面很难给这个产业划定清晰的边界，它包括整个供应网络、开发创新产品的研究机构、技术设备的制造商、电信服务的提供商、设备和程序的销售商以及软件开发商等。由于电信业与计算机、娱乐等产业的融合，产业边界已变得模糊不清。

需求方面的边界也同样被打破。在过去，你可以很清楚地区分人们对于电话、邮件、电报、图片摄影、学习和各类特定信息服务的不同需求。但如今，任何一条科技路线都能满足多方面的需求。边界的渗透让我们不可能明确界定哪些“供应群”与哪些“需求群”相对应。

任何一家企业都适用规模报酬递增的规律。产业进入需要的初始投资和固定成本规模巨大。销售额越大，固定成本的分摊面就越广。因此，所有的企业都在努力游说，让尽可能多的客户购买自己的设备与服务。企业收入的增长仅仅受制于销量的多少，而不是出于成本方面的考虑。

熊彼特式的“创造性破坏”的进程不断上演，创造的速度超越了破坏。鼓励产业进入和扩张的力量极其强大，而产业退出和淘汰却存在强大阻力。已经进入产业的厂商即便出现亏损，也会竭尽全力维护自己的地位。例如，随着移动电话的普及，固定电话服务受到排挤（图 1），但其使用者的流失速度非常慢，与移动电话用户的扩张速度完全不成正比。增加一名手机用户，并不代表减少一名固定电话用户。这两个电信网络已经并存了相当长的时间，把二者打通的技术手段正在研发之中。各种类型的电信技术及其衍生发展，制造了规模极大的过剩产能。或者说，在电信服务业存在明显的过度供给现象。

这不过是几十年时间里在电信业所发生的产业变化。如果市场研究者们曾试图在这个漫长历程中任一时点采访生产商和服务商的负责人，询问对方能否为远超过实际客户数的对象提供电信服务，答案恐怕是由衷的肯定。这是一种

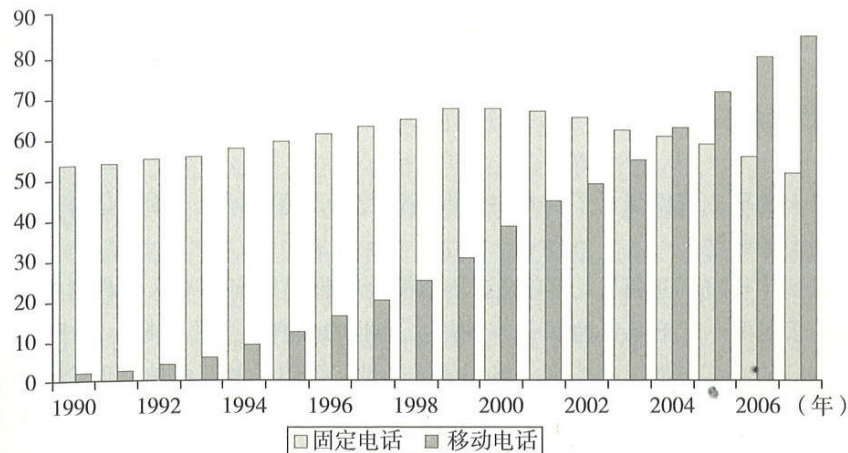


图 1 美国每 100 名居民拥有的固定电话和移动电话数 (1990 ~ 2007 年)

资料来源: International Telecommunications Union (2007)。

典型的过度供给症状，我们可以认为电信市场处于长期的过度供给状态。

接下来，我们再对比苏联和东欧社会主义国家在 1980 年代的情况（表 1）。当时的很多居民长期装不上电话，要么需要数年时间排队等候安装，要么不得不利用其他门路加塞。成功装上电话的人则发现，打电话是件恼火的事情，线路经常过于繁忙，接通需要很耐心的等待。如果有人向国有电信企业的负责人提出同样的问题：你们有没有能力给更多的用户提供服务？回答肯定是响亮的“没有”。他们还会特别提醒说，排队等着装电话的人多着呢！在社会主义经济制度下，电信业处于长期的过度需求状态。

表 1 部分社会主义国家和美国每 100 名居民的固定电话用户数 (1980 ~ 1988 年)

国家	1980	1984	1988
保加利亚	10.2	15.6	21.3
匈牙利	5.8	6.6	8.2
波兰	5.5	6.4	7.8
罗马尼亚	7.3	8.4	9.8
前苏联	7.0	9.6	12.3
美国	40.8	46.8	50.8

资料来源: United Nations Statistics Division (2009a)。

于是我们看到，电信业在一种经济制度下表现为过剩经济，在另一种经济制度下表现为短缺经济。<sup>①</sup>

## 供给形成过程

以上的历史案例显示了资本主义市场运行的普遍特征，下面的任务则是进行理论上的归纳。

提到“市场的运行”，所有经济学家都会联想起任何一本入门教材中都会出现的著名图示：产品的边际成本曲线、企业的利润最大化目标、需求曲线，以及它们的交汇所形成的均衡价格和最优产量。该图示及其他抽象理论所描述的是企业从财务角度进行决策的静态场景，但我们这里要将其超越，我们需要突破经典模型的大多数严格约束及其静态特征，关注市场的动态发展，为此不得不采取没有那么精细的描述式的论证。

首先从供给开始，其路径是个动态过程，在以下几方面发生变化：

- 市场的参与者在持续发生改变，包括提供产品和服务的单位和企业的集合以及每个生产商和供应商：不断有新企业的进入、现有企业的合并与分拆以及退出等。这个集合的构成呈现经常的变动。<sup>②</sup>

- 可供买方选择的产品和服务的具体种类在持续发生改变。随着产品的多样化发展，选择空间从长期来看在不断扩大。新产品陆续面世，而老产品逐渐消失。从个别生产商和供应商那里就能看到产品种类的不停变化，如果从企业群体、整个产业乃至整个国家经济的角度来看，这种现象将更加明显。

- 在产业部门或宏观经济层面把数据汇总起来，会发现短期波动和长期波动持续存在。从历史角度来看，还存在长期的持续增长。

随着产品种类的不断改变，持续的过剩将会出现：没有被买家采购的产品，或者没有被厂商利用的产能。这些滞销的存货与闲置的产能的形成同样是一个动态的过程。

过剩是如何形成的？过度供给的现象为何会普遍出现？有四方面的作用可以为此提供解释。

<sup>①</sup> 为保持客观，我需要补充指出，制度的改变与技术的飞跃同时发生，特别是移动电话的引入。即使没有技术的飞跃，制度改变也可能促进电话服务的普及，但程度如何难以判断。

<sup>②</sup> 参见 Bartelsman 等人（2004），该研究包含 OECD 各国的详细的企业组成状况。

（1）垄断竞争。在张伯伦 [Chamberlin, 1962 (1933)] 和罗宾逊 [Robinson, 1969 (1933)] 的开创研究之后，不完全竞争理论已被主流经济学接纳。这一重要理论证明垄断竞争企业存在过剩产能，为我的观点提供了佐证。在确定产量水平的时候，垄断竞争企业的行为与完全竞争市场上的表现不同，为实现利润最大化的目标，它们不会充分动用自己的产能。<sup>①</sup>

我们这里关注的是市场结构不完美时的表现，其正式的说法是“垄断竞争”（采用“寡头”一词可能更准确，但研究文献里采用“垄断”的居多）。<sup>②</sup>

参与垄断竞争的企业在与对手争夺客户的时候不但可以通过价格优势，也能利用与价格无关的其他条件：更好的质量、更殷勤的服务以及送货上门等额外的服务等。最重要的是可以引进新产品，这个问题还将做单独考察。<sup>③</sup>

（2）需求的不确定性。让我们看一个零售商店的案例，这是家常见的出售某些特定种类产品的商店，如家用电器和其他电子产品，包括电冰箱、笔记本电脑、电视机以及打印机电缆线等。这是家零售店，过剩现象以存货的形式出现。但我想指出的是，这并不限于零售业，也适用于所有产品和服务的生产。我们的分析并不只针对产品库存，也包括那些能够立刻或在很短时期内提供产品和服务的过剩产能。换句话说，采用零售业的案例不过是为了更直观起见。

再回到这个案例上。该商店无法准确预测有多少客户来光顾以及客人们都要寻找何种产品。但店主希望，尽量不要让客户空手而去。我们把得到满足的客户占所有来访者的比例称为该商店的“安全水平”，<sup>④</sup> 假设店主已知哪些种类的产品会对来访者产生吸引力，但不清楚需求的总量和具体构成，那么很明显，店主希望实现的安全水平越高，他需要保持的存货量就越大。

<sup>①</sup> 多马的一篇杰出论文用图示描述了垄断竞争的这个特点，与本文的主题有关：为什么资本主义市场经济会系统性地出现过度供给现象 (Evsey D. Domar, 1989)。我发现许多研究成果都在某些方面和我自己的研究思路有类似之处，并能够借用一二，但多马的论文是唯一一篇让我感觉到完全相通的，而不仅是细节。

<sup>②</sup> 我不知道有什么统计数据可以显示，某个国家在某段时期内有多大比例的商品和服务交易是发生在完全竞争类型的市场中，我只能猜测，这个比例可能相对比较小，垄断竞争应该居于统治地位。

<sup>③</sup> 魏茨曼的模型 (Weitzman, 2000) 对社会主义和资本主义经济进行了比较，发挥核心作用的是变量 E，指“销售努力” (selling effort)。魏茨曼的研究显示，垄断竞争要求生产商和销售商做出较高的销售努力，而计划经济中的 E 值较低。

<sup>④</sup> 或许可以用其他办法来表示“安全水平”，这取决于测算的时期长短，例如，感到满意的购买量占全部意向购买量的比例，而不是满意的人数与全部客户人数之比。

有关企业运营的研究提供了几种从数量上分析这一问题的模型。<sup>①</sup> 商店所需要的库存规模受多种因素的影响, 例如:

- 尽管需求存在不确定性, 但优秀的卖家仍然知道其大致构成、规模和波动幅度 (Ramey 和 West, 1999)。

- 存货通常采取分批式和突进式的办法更换, 而不是陆续的更换。卖家所需要的存货水平往往是与供应商共同商定。

- 每个空手而去的客户都意味着损失, 但堆积时间太久的存货也是浪费。这两种成本需要仔细权衡。

本文是针对经济学理论的研究, 而非存货管理。不过, 从零售业的实践中还是可以看到, 有些商店在订货和库存管理上比同行更有效率。许多生产商和销售商完全没有精心制定的存货策略, 纯粹是依靠经验产生的直觉。然而对于一家在严酷的垄断竞争市场上打拼、试图维持和提高市场地位的销售商来说, 最主要的武器之一仍然是保持较高的安全水平, 尽量满足客户的需求, 而这可以通过规模较大且精心挑选的存货来实现。

在较高的安全水平上满足客户需求, 这是过剩经济的一个主要特征。由于每家商店自身的利益, 总体上出现了过度供给的状况, 这指的是某个行业内所有卖家的总和。

除零售业以外, 这一理论对生产企业面临的存货与产能储备问题同样适用。

传统的企业经营模型研究者希望找到适用的商业策略, 他们可以将安全水平视为给定, 然后推导出与之匹配的库存和产能储备的规模和构成。而本文采取的实证研究方法的思路相反: 在库存和产能储备给定后, 推导出能够在多高的安全水平上满足需求? 以前还从来没有人对这些指标在资本主义或社会主义经济制度下的不同表现进行观察和测算, 实在令人汗颜。但我相信, 安全水平在过剩经济中较高, 而在短缺经济中较低。<sup>②</sup>

<sup>①</sup> 有关存货和产能储备的研究文献很丰富, 涉及多个学科, 特别是微观经济学、企业运营和管理科学等 (如 Chopra 和 Meindl, 2003, 第 11~12 章; Toomey, 2000)。这些数学模型采用了不同的方法, 如随机过程理论和随机程序理论等 (Prékopa, 1995)。

<sup>②</sup> 在短缺经济中, 由于购买方的普遍接受, 差距被缩小, 这是因为几乎不可能在商店里找到某些需要的商品。这个现象类似于资本主义国家的劳动力市场上因为长期沮丧而放弃寻找工作的人。

(3) 创新与创造性破坏。在技术进步陷入停滞时, 导致过剩现象不断累积的力量同样会发生作用, 但技术并不会停滞, 而是不可遏制地向前发展。<sup>①</sup> 在垄断竞争市场中, 如果某个生产商或销售商能为客户提供其他厂商所没有的新产品或新服务, 将获得巨大的竞争优势。这种描述更适用于那些在世界范围内引入革命性的新产品的创新者, 但也符合被先驱者唤醒、在本国或其他国家快速引入革新的追随者的情形。

创新会如何影响经济中的过剩现象呢? 某个地方出现了某种新事物, 并开始吸引需求, 并不代表老产品会马上消失, 还有许多因素支持它们继续存在。

工厂中使用的某些工具和经验是专门针对某些产品的, 创新的引入需要进行资本投资、重新培训员工或招募新的劳动力。如果不能适应变化, 以前的所有投资都将遭受损失, 不但包括工厂, 还有引进和推广所花费的体力与智力投入。生产老产品的工厂已投入了大量资本和劳动力, 雇主很自然急于获得回报, 雇员们也希望保住饭碗。很多时候, 政府还会帮助那些面临淘汰的企业继续存活下去 (这种现象我称之为软预算约束, 将在下文专门讨论)。

以上因素最终意味着, 在现有的产能之外, 将会不但出现生产或提供新产品或新服务的产能。尽管老产能的比重会缓慢下降, 新老产能之和却会超出总需求。这正是出现经常性的过度供给状态的主要机制之一。

借用熊彼特的名言来说 [Schumpeter, 2010 (1942)], 创造的速度超过了破坏的速度。在偶然情况下, 即便不存在这一速度差, 过剩仍然会不断产生。如果一开始存在过剩, 即使创造的速度与破坏的速度相当, 初始的过剩比率 (如过剩产能与总产能之比) 依然会持续存在。这个变化要想反方向发生, 即产能淘汰的破坏速度要在比较长的时期里快于创造的速度, 却是非常不容易。我还没见过哪种老产品的消失会快于新产品的推广。

以上的概述可以用规范的理论模型来阐述。“创造”与“破坏”的过程可以用统计上能够观测的指标来表示, 两个变化方向都存在可以证实或证伪的命题。

请注意, 以上三种作用背后的动力机制是相同的: 卖方之间的竞争与过剩现象之间的联系。过剩既是竞争的原因, 又是其结果。如果在某个市

<sup>①</sup> 我的文章《创新和活力》(Innovation and dynamism) 总结了这方面的结论和文献综述 (Kornai, 2010), 其中的论点和本文密切相关。

场上，需求恒等于供给，卖方何来竞争动力呢？那将陷入一种麻木的静止状态。正是由于过剩的存在，让买方有所选择，竞争才成为可能，才会被激发，因为卖方都希望消除过剩。反过来，竞争与对抗又在持续地创造新的过剩。

(4) 规模经济。在标准的微观经济学（和年轻经济学家们深信不疑）的简化模型中，生产商的平均成本曲线都被假定为“U”形。产量超过曲线最低谷之后，边际成本会上升。因此，企业不愿意让产量超过某个水平。这不是外加的约束，而是企业自身出于利润最大化的考虑而不愿意再增加产量。

在边际成本随产量的增加而下降的时候，情况有所不同，或者说，出现了规模报酬递增。后者在实际经济生活中非常普遍，当生产商、服务商或零售商的固定成本在全部成本中占据较大比重时就会出现。产量越高，单位产品分摊的成本就越低。此时不存在对提高产量的内部约束，只有外部约束发挥作用，主要是指需求量的限制。企业在不得不停止增加产量时，并非是出于自身的成本考虑。

规模经济会在两个时间维度上导致过剩产能的出现：长期和短期。面临垄断竞争的大企业对新产能大胆（甚至莽撞）投资，主要的动力是它们认为工厂越大，规模经济效益就越显著。尽管整个产业可能因产能的闲置而受损，新的大型工厂仍然会持续建成。这方面的典型案例是汽车产业，下文将进行讨论。

在短期内，大多数面临规模收益递增的企业都没办法把产能利用到最大限度，虽然这样做最有利，但通常来说需求方面的约束会极大地限制产量。于是，决策者们能切身感觉到没有利用的产能的确是“过剩”的产能，企业的利益要求他们尽可能地增加产量，同时也有能力增加产量。在规模收益递减的时候，决策者对形势则有不同的判断。虽然从物质和技术上说，他们也可以增加产量，但利润却会减少，因此在达到利润最大化的产量后会毫不犹豫地停止增产。<sup>①</sup>

以上，我分别介绍了四种不同的作用力。在本文中，它们的共同点是对过

<sup>①</sup> 规模收益递增的重要性受到很多经济学家的重视，尤其是考尔多（Kaldor, 1981）和阿瑟（Arthur, 1994）。规模收益递增在几个方面产生影响，我只是提到了过剩产能的方面。另外一本名著（Helpman 和 Krugman, 1985）则分析了它对于国际分工和国际贸易的影响。

剩经济的影响，导致了长期的过度供给现象的再现。很多时候，我们会看到两种、三种甚至四种作用同时出现，互相影响和加强，对供给过程产生合力，推动过度供给的形成。

在观察其中某种作用时，大多数研究者会假设企业的目标是利润最大化。这对设计数学模型较为方便，主流经济学家们还到处宣传这个观点。但更全面的社会学和社会心理学研究表明，把“利润最大化”作为企业行为的主要特征并非普遍适用。“公司”不是一个毫无人格特征的无形机构，其决策是人制定的。在现代的大企业的决策中，老板（包括比较有发言权的重要股东）和高层经理人都发挥重要作用。他们的利益经常存在冲突，决策结果取决于彼此间的相对力量的大小及达成怎样的妥协。时间因素也需要考虑：他们争取眼前利益的愿望有多强烈，又在多大程度上看重长期收益。

决策者的动机不能用“利润”一词来简单描述，也来自对权力、威望和改善公众形象的渴望，尤其是成为市场份额最大占有者可能带来的尊重和荣耀。<sup>①</sup> 这些动机往往是一致的，但也可能发生冲突。如果企业的决策者并不是在追求利润最大化，或者不仅仅是追求利润最大化，也受到其他动机的影响，上文描述的四种作用也同样适用。奋力拼搏的企业家们眼前看到的不见得只有利润最大化，竞争精神、争当第一的愿望，也能产生巨大的推动力。同样重要的还有追逐权力带来的扩张冲动：“让我们成为最大、最强！让我们来统治市场！”除其他动机外，企业家们的行为还来自本能、生命力以及凯恩斯所说的动物精神等 [Keynes, 1967 (1936), 第 161 ~ 162 页]；Akerlof 和 Shiller, 2009]。在我的《短缺经济学》一书中，这种动力被称为扩张冲动（Kornai, 1980）。

企业决策者们受到各种行为模式的影响，每个决策者的组合倾向都不相同。<sup>②</sup> 但即使决策者不考虑利润最大化或不仅是考虑利润最大化，也受到其他驱动因素的影响，以上四种作用机制仍然适用。

<sup>①</sup> 行为经济学的许多著作分析了经理人的行为背后的动机。除各种经济心理研究以外，还有大量研究各种动机的心理学理论。经济学家们还没有把这些新的心理学发现充分利用起来。

<sup>②</sup> 很难找到一群人中的“平均值”，因为那样会否认个体的异质性。这对心理学家甚至作家来说或许不算什么，但经济学家们却需要慎重对待，用批判的眼光来审视那些忽略决策群体的异质性的模型（参见 Kirman, 1992）。

## 需求形成进程

需求的形成同样是个动态过程，不但受买方的偏好、收入或财富的影响，也取决于任何时点的供给等因素，这一点对目前讨论的话题尤其重要。例如，新产品的出现会唤起新的需求，让老产品显得过时。虽然主流经济学家们深信不疑，实际上却不存在交汇于某点的两条独立曲线，因为供给在任何时候都是需求的主要解释变量之一，反之亦然。

供给和需求的形成过程相互影响，如果它们的变化趋势基本保持平行，以相似的速度增长，其交互作用依然存在。不过，本文的目的是解释为什么在过剩经济中需求会落后于供给，在需求得到基本满足后，还会有多余的供给出现。

上节介绍了抬升供给过程的一些作用力，下面则讨论打压需求过程、制约需求扩张的某些作用。

其中最重要的是雇主与雇员之间的利益冲突。抵制雇员的加薪要求符合雇主的利益。对劳动的报酬不能超过工作的边际收益，这在雇主们的考虑之中。<sup>①</sup> 如果雇主能够把工资压低到该水平之下（或许可以通过长期失业来实现），就能进一步节省成本。反之，如果雇员的加薪要求实现，则会侵蚀企业利润。但无论争夺的结果如何，在资本主义经济制度下，工资的任何膨胀都会受到雇主利益的反向制约，最终导致雇员的收入水平会制约产量的扩张。

随着产量和生产率的长期增长，雇员的收入也会提高。但工资具有粘性，不能很密切地反映产量和生产率的增长。于是，我们会看到需求落后于生产的积极扩张的趋势。

这个观点与马克思主义有着密切联系（《资本论》，尤其是第23章和第25章）。在许多重要问题上，我并不赞同马克思主义政治经济学的主张，但我承认在这个论点上存在相似之处。<sup>②</sup> 资本主义制度下的生产带来的供给扩张并不

<sup>①</sup> 或许可以支付更高水平的工资，例如对某些具有特殊技能或忠诚的雇员群体支付所谓的效率工资。本文将在讨论劳动力市场那章再进行讨论，这里只是强调即使效率工资也要物有所值，不能上涨过高。

<sup>②</sup> 我在以前的多篇文章中做过详细阐述（包括 Kornai, 2008）：我在哪些问题上接受马克思主义的研究方法，针对某些现象在多大程度上应用这些方法，以及在哪些问题上不赞成马克思主义的主张。

是由于政府的财政和货币政策，尽管这些政策可以有效地促进或制约供给水平。扩张的刺激来自企业所有者与经理人的固有利益，这种固有利益同时又给需求的增长设置了障碍。导致过剩的驱动因素是在微观层面上产生的。<sup>①</sup>

以上介绍的是利益冲突机制造成的影响，如果家庭和企业的支出受到的预算约束较为严厉，效果将更加显著，它会限制过度增加支出和放松工资约束。

预算约束的软硬程度从看似矛盾的方向影响着供给的形成过程。一方面，企业和家庭的预算约束必须足够硬，以防止需求的过度膨胀。另一方面，人为地保护落后企业、放松预算约束，又会继续增加过剩产能。不是因为我的描述和解释的逻辑出了问题，才导致了相反的结果，而是客观上存在着两种相反的影响。

## 定价过程

根据正统的新古典经济学理论，价格是解决供需失衡的必要手段。出现供给过剩的时候，价格必然下跌，从而提振需求。价格变化的确会在一定程度上发挥这个功能，但仅仅是在一定程度上，这不但是由于随机误差的影响，还源自系统性的扭曲。

定价本身也是个动态过程，不能像瓦尔拉斯的试探理论描述的那样，被视为向均衡价格水平趋近连续过程。在价格到达均衡水平前，供给和需求都已经改变了，我已经指出，这种变化是在不断发生。所以并不存在向某个特定目标的趋近，目标本身就在变动中。

人们现在普遍认为，资本主义市场中的价格具有粘性，一些较为明智的主流经济学家（Mankiw, 1985；Blinder 等人，1998，Bils-Klenow, 2004）和有的非主流学者，如后凯恩斯主义（Lee, 1998）都能赞同。价格粘性有很多原因：对供需关系的变化存在滞后反应，企业不愿意承担调价带来的附加成本（所谓“菜单成本”）等。

垄断竞争市场中的价格由卖方决定，买方要么接受，要么转向其他卖方。<sup>②</sup> 价格粘性具有不对称的特点：下调更为困难。即使感受到供给过剩，卖

<sup>①</sup> 其他人的论述中有类似的分析（Bhaduri, 2007）。

<sup>②</sup> 价格制定者和价格接受者的划分是西托夫斯基引入的（Tibor Scitovsky, 1985），他特别指出，垄断竞争市场的卖方制定价格，造成了多种影响，包括闲置产能和存货的堆积。



方仍然担心利润下跌，而不愿意永久性地降价。<sup>①</sup> 所有企业（和货币政策制定者）都会对通货紧缩及其对宏观经济的破坏感到恐惧。我希望强调价格粘性的这个不对称性，因为它是资本主义市场经济普遍的不对称性的一个主要原因，这个普遍的不对称性就是过度供给占据主导的倾向。

在供给、需求和价格的共同运动中，价格变化不但不能消除普遍存在的过度供给现象，还会继续诱发其产生。价格波动往往局限在过剩经济的典型范围内。

### 3. 产品和服务市场：概念工具和测算方法

我现在兑现之前的承诺，开始介绍概念工具，讨论对现象的观察及其测算方法。

供给和需求到底是指什么？那些曾毫不犹豫地接受这些概念的人，或许可以不要那么自信。<sup>②</sup> “过度”一词也已多次出现，但我还没有准确描述是什么东西过度，过度供给是相对于什么标准而言？

#### 容易理解的假想案例

(1) 一家处于过度供给状态的服务企业。首先请设想，我们现在去采访一家服务企业（例如酒店）的经理人。<sup>③</sup> 这家酒店是在过剩经济中经营。酒店业习惯采用的经营指标是“客房入住率”或其对立指标“空房率”，均以百分比表示。我们要问这位经理，“您下周能够实现和希望实现的客房入住率有多高？”下面是几个附加条件：

- 下周的客房价格由酒店自己决定。
- 问题仅针对一个很短的时期，酒店在短期内的接待能力是固定的，由可

① 比较容易通过临时降价和促销来处理存货。

② 我曾经在《反均衡》(Anti-Equilibrium, Kornai, 1971) 一书中提到这些被广泛接受和采用的术语的问题，指出其定义含糊不清。我的思考没有激发出太多反应。在40年之后回顾当年的写作，我认为自己的看法依然是正确的。结合后来增加的知识，我在这里再次表述自己的批评意见以及在概念和方法论上的改进建议，希望这次能获得更大的反响。

③ 本文有关供给的讨论都是针对企业，但观点也同样适用于个体的生产商、贸易商和服务商。

用的客房数目决定。仅从实物角度考虑会得到一个理论上的产能，但经常还要克服其他障碍，如维护和技术等问题，会导致某些房间无法提供服务。考虑到这些后，经理人会给出一个现实的产能。<sup>①</sup> 在这个案例及下面的研究中，如果现实产能被百分之百地利用，则视为全部利用。<sup>②</sup> 回答者会根据对现实产能的冷静分析进行回答，而我们感兴趣的也是提供服务的实际意愿，而非梦想或奢望。

• 不要求对方预测可能实现的业绩，或者说希望来多少客人。让他假定找房间的客人会在任何一天的任何时候出现，或者说，他要考虑的是，仅从酒店的角度来看，愿意接待多少客人，希望实现多高的入住率。

我当然希望酒店爆满，把现实的接待能力都充分利用起来——如果对方不这样回答，那才出人意料。于是在下周，我们会得到酒店的实际入住率，假定其平均水平为75%，显然，对过度供给的量化描述应该是25%。

我们假定，对某个城市的所有酒店都进行这样的采访，并收集实际入住率，看到的都是过度供给状态。然后，我们就能得到对这个局部市场的供需情况的准确描述。把酒店经理人的预先测算汇总起来，即能得到某个星期的供给量，把事后的实际入住率汇总起来，即能得到需求量。在这个案例中，总供给显然超过了总需求。我们不需要向顾客询问需求水平，因为普遍存在的过度供给表明，需求量等于实际交易量。

这一结果可以称为“短缺方法则” (rule of the shorter side)。假如存在过度供给，需求就处于短缺方，实际交易水平始终等于短缺方的水平。

这个法则对观察和测算方法具有重要影响。从实际交易和销售数据中，只能估计出短缺方的规模。对过剩方的测算需要用其他方法。在我们设想的案例中，采取了最简单的方法来估算过剩方：询问酒店负责人希望销售的房间数量。如果考虑到酒店希望把所有房间都销售出去，所提的问题或许应该改为请

① 美联储会定期公布产能利用指标，其定义是：“美联储编制的产能利用指数采用的是可持续的最大产量 (sustainable maximum output) 的概念，指一家工厂根据现实的工作计划、考虑到正常的检修期、在劳动力和原材料供应充足的情况下可以实现的最大产量。这一概念基本对应生产函数的最大投入点，代表可持续的最大产出，而非短期内的不可持续的更大值。例如，某家企业可以推迟日常的维护，或者在短期内超负荷生产，但从长期而言是不可持续的” (Morin - Stevens, 2004, 第3~4页)。

② 尽管实际产能通常低于100%，还是可以假设有超过100%的产能的情况。例如在这个案例中，为应付太多的客人，酒店可以推迟日常维护工作。

他们预先测算实际的供应能力。

回到之前的话题，“过度”的定义涉及怎样的参照标准。这个案例给出了清晰的回答：供给超过了实际交易量，而实际交易量等于需求。

(2) 在过度供给状态下，一家销售库存商品的贸易商。<sup>①</sup>零售活动可能受到多种因素的限制（如单位的物质条件或员工的生产效率等），但供给的主要约束还是现有的库存。我们要测算一段时期内（如一周时间）的供给水平，而不是限于某个时点，这和上面的服务企业的情况相似。我们向这家贸易商的经理人所提的问题是：在普通的价格水平上，贵公司能够并愿意向客户销售多少产品，包括期初的存货和这一周内新增的存货。可以预想，回答应该是“全部存货”。<sup>②</sup>贸易商能够并愿意销售的数量可以视为市场的供给。

为了得出清楚的结论，还需要加上如下两个条件：

- 要新增的库存在采访前已经预订。我们要求回答者假定，在未来一周不再调整存货订单，维持原来的对于“短期”的解释。

- 第二个条件针对我们——思想试验的设计者。我们假定，所有产品到这一周的周末都还有存货，没有产品被卖光，或者说，所有产品的需求均不足。

如果满足以上两个条件，到这个周末的时候，我们就能准确测算出过度供给的规模有多大，它就等于期末的存货。这个结果可以用简单的等式来表达：

供给 = 期初存货 + 新增存货

需求 = 实际交易或销售额（短缺方）

过度供给 = 期末存货

期末存货是卖家事前希望卖出去但未能实现的部分，是事后得到的结果。

(3) 处于过度供给状态的一家生产企业。我们再进行一次假想采访，这次的对象是生产某种有形的耐久产品的企业（如汽车厂）的经理人，我们已知采访是在过度供给的市场中进行。

如何设计问题才能发现供给量，这次比前两个案例要困难。与前两次采访相同的是，我们的问题仅涉及下一周这个很短的期限，我们询问的对象是销售人员。问题、假设和条件与贸易商的案例基本相同，我们要问，这家汽车厂根据现有的成品和未来一周将完工的汽车数量，能够并愿意销售多少辆汽车？注

<sup>①</sup> 贸易商可以发挥服务中介的作用，本文未考虑这种特殊情况。

<sup>②</sup> 商店的全部存货到这一周末被卖光显得有些夸张，就像本文开头介绍的华沙的食品店那样，那会严重挫伤客户。但这种情况在市场经济中很难发生，因此可以忽略不计。

意：成品的完成速度不再改变。于是我们能得到下周可用于销售的总的供给量。过度供给的规模（未能售出的汽车数量）将等于期末的存货。

如果接受采访的经理人是负责安排生产的，我们会根据完全不同的假设提出不同的问题。对于汽车这样复杂的产品而言，在一周这样短的时期内进行估计是不合适、不符合习惯的，因此我们要问这位经理人对下个季度的看法。

这次采访及其背后的条件与第1个案例（服务企业）类似，受访者需要假设有足够多的订货，因此，愿意生产多少产品完全取决于他们自己。答案会尽可能接近这家工厂的实际生产能力。如果产量在接近实际产能的时候，满足规模收益不变的条件，那将很容易得到这样的回答。如果出现规模收益递增的情况（对汽车厂而言往往如此），则更是如此。

后面的结论也和服务企业的案例相似。如果收到的订单太少，最后的产量就会小于实际的产能，或者由于销售状况不能得到改善，产量目标就会被下调。于是在季度末，我们会看到企业未能充分利用自己的全部产能。事后得出的闲置产能的规模可以被视为过度供给。

请注意，生产企业面临两种类型的过度供给：希望出售却未能实现的存货，以及希望利用却未能实现的产能。两者都是指汽车，但这两个确定的汽车数量却不能加总。前者是场地里停放的现实的汽车，后者则仅限于想象，是对没有实现的生产愿望的量化描述。从不同视角看，未能出售的存货与未能利用的产能都可以视为过度供给，但把它们加起来就不太好解释了。

以上是通过假想的采访对有关概念进行描述。幸运的是，现实生活中有人就类似问题进行过真实的采访。

表2是对工业产能利用水平的国际比较。研究者分析了34个国家从1978~2008年的数据，表2的数据来源包括13个国家和超过100个观测值，具体的数据是每个国家在某个时期的中值。

表2 产能利用率的国际比较（1978~2008年）

国家	中值	观察值数量	标准差
比利时	79.0	121	2.89
加拿大	81.3	152	4.11
法国	84.4	130	2.02

续表

国家	中值	观察值数量	标准差
德国	83.6	154	3.51
意大利	75.7	154	2.70
日本	79.0	153	8.11
荷兰	82.5	147	2.58
新西兰	89.2	153	2.13
挪威	82.4	138	2.70
葡萄牙	78.9	126	2.49
西班牙	79.8	154	3.03
瑞士	83.8	154	3.33
美国	80.4	145	3.91

注：数据显示的是制造业的产能利用率，“观测值数量”一栏表示某个国家在一个季度内获得的数据的数目。

资料来源：Etter-Graff-Müller (2008, 第8页)。

下面的图2反映的则是美国的产能利用水平的时序指标。该报告的作者特别指出，实际利用水平是与长期的可持续生产水平（即实际产能）进行对比（也可参见 Corrado 和 Matthey, 1997）。

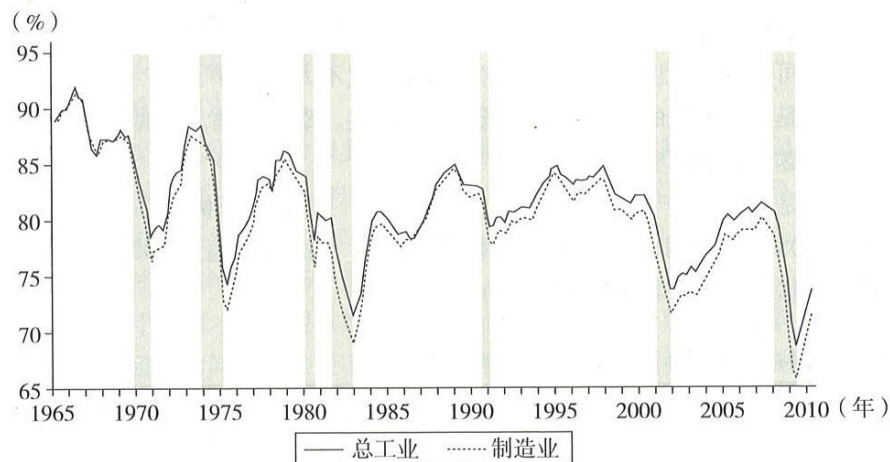


图2 美国的产能利用率（1965~2010年）

注：灰色的柱状区域代表衰退期。衰退期根据 NBER 的定义确定。

资料来源：Federal Reserve Statistical Release (2010)。

另一个案例是法国的时间序列图（见图3）。

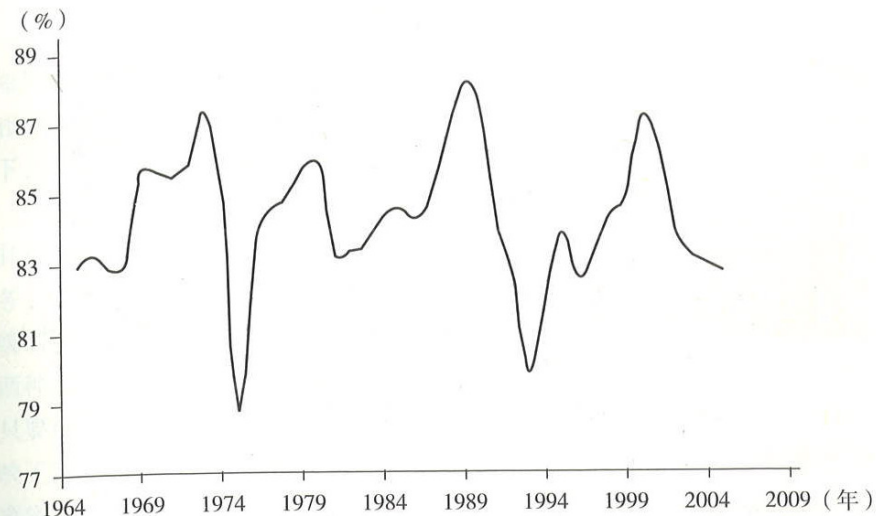


图3 法国的产能利用率（1965~2005年）

资料来源：Allain-Canry (2008)。Olivier Allain 与 Nicolas Canry 提供了数据资料，对此我深表感谢。

从这些图表中可以清晰明快地得出如下结论：

各国的产能利用率不同，年度间的利用率有波动，但从未接近90%的水平。这支持资本主义经济制度下存在普遍、长期的大规模闲置产能的观点。

表2和图2、图3显示的是工业产能的利用率。有人正对住房存量的利用率做类似的调查。从本文的研究目的出发，业主自有住房和出租住房的出售和出租数据可以加总：其总数可视为住房产业的未利用产能。图4明确显示，美国在任一年份的这一加总比率从未低于8%，在2004年更升至12%，这同样意味着存在相当大比例的过度供给。

### 第一个难点——供给和需求的持续的共同调整

以上介绍的三个简化案例提供了“过度供给”的概念定义及其测算方法，当然我有意识地对此进行了各种抽象，以便尽快审视测算的结果。接下来，我希望与读者一起研究实际情况，分析真实的商业活动的情形，它们与上面三个

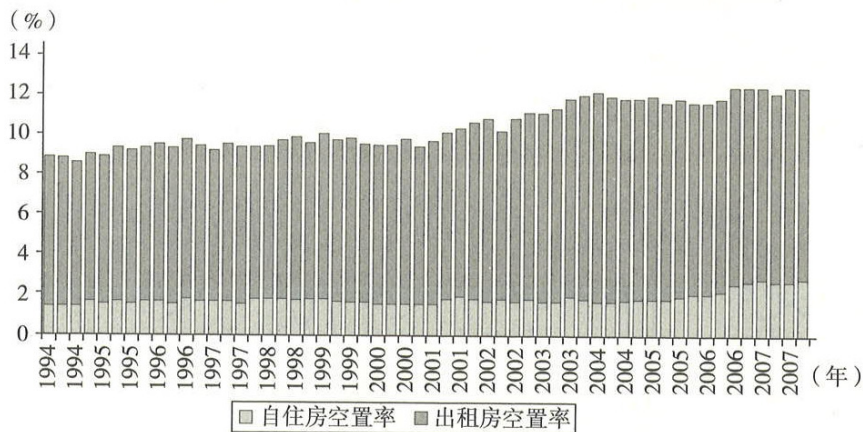


图4 美国的出租房和自住房空置率 (1994~2007年)

资料来源: U. S. Census Bureau (2010)。

理想化的简单案例存在很大不同。

我不是担心观察者或统计者可能犯的错误及其引起的随机误差。经常遇到的一个问题是:在某些情况下,不可能“客观”地坚持采用以上的定义和测算方法。

这样一个难点已经在前面的第2章中有所触及,描述需求的形成过程。我们现在再从测算的角度来同时观察供求两方面的变化。

对每个卖方和买方而言,供给和需求两方面的变化共同构成了一个动态过程。生产商和贸易商的销售倾向在不停改变,针对现实情况不得不出调整,这个过程哪个时刻才代表“真实”的供给呢?市场研究者们往往会问生产企业的经理人计划生产或销售何种产品,然后把计划与实际的生产与销售进行对比。实际结果可能大于或小于计划。<sup>①</sup>问题在于,我们要搞清楚经理人在接受提问的时候正处于决策形成过程的哪个阶段。有关的需求信息会怎样影响供给形成的过程?经理人越是靠近供给对需求做出必然调整的最后时点,他制定的计划与实际产出就越是一致。

买方的购买愿望符合同样的逻辑。这样的愿望会在哪个时点转变为“真实”的需求?客户在什么时候开始进行购买?在什么时候进入商店?在什么

<sup>①</sup> 例如广泛使用的商业调查数据 (Business Tendency Survey, OECD, 2003)。

时间发现没有找到预定的物品,转而购买替代品?或者在不久之后,会在什么时候去第二家或者第三家商店购买其他商品?

我们可以追踪每个卖家或买家的这些决策过程,但如果希望将某个时点或短期内的大量买方或卖方的购买或销售意愿加总起来,就会把不同性质的成分相加。<sup>①</sup>其结果导致,总数的概念会变得模糊不清,在没有更多信息的情况下,不可能做出合理解释。

我们利用酒店业的案例做一番详细阐述。如果某个城市来了20位客户,其中10位寻找四星级酒店,另10位寻找二星级酒店,他们都如愿获得了服务。但如果所有20位客户都想找四星级酒店,而其中10位最后只找到二星级酒店,那么,就只有10位客户的需求得到满足,另10位感到失望。但在这两种情况下,酒店业的入住率这一实际统计指标是相同的。不管是什么情况,只要每位客户都能找到房间入住,按“短缺方法则”,实际的房间入住数都反映了客户的总需求。然而,这一总数掩盖了某些客户未能获得理想服务的事实。

## 第二个难点: 过度供给与过度需求同时发生

在本章开头的讨论中我进行了简化,假设在每种交易中需求都是短缺方,排除了过度需求的可能性。结合“短缺方法则”,很明显,在这样的案例中那些描述性的统计数据反映的是需求量,因为我们已知是在讨论过度供给的市场。

然而在现实生活中,我们却不能如此肯定,可能同时存在过度供给与过度需求的现象。某些酒店接待的客户可能很少,而有的客户同时又没能找到自己喜欢的酒店。还有,每天都在发生调整。例如,我们可以编制覆盖某个城市的所有酒店的入住率的数据表,但仅靠这个报告并不能证明供给能否满足需求。

过度供给与过度需求的现象不但可以并存,还会发生相互作用。某个拥有过剩产品的卖方可能缺乏资金,从而无法购买自己需要的产品,他的需求因此下降,导致另一位卖家的供给过剩,形成连锁反应 (Clower, 1965, 1967; Leijonhufvud, 1968)。因此,过度供给现象会产生溢出的乘数效应。

<sup>①</sup> 异质性问题在前文也提到过,它给供给和需求的记录带来了严重的困难。

如果微观上的过度需求和过度供给数据能被分别统计,那会是件大好事,可以对这些现象的分布进行描述,分析其随机特征。

### 另一个调查:观察生产的障碍及其微观约束

前一小节分析了过度供给与过度需求并存的问题。下面介绍的规范国际比较可以有助于对它们的理解和观察,但需要正确解释采用的方法。

数十年以来,每个季度,26个欧洲国家的26家国际商业研究机构都会对若干工业企业的某些经理人代表提出以下问题:<sup>①</sup>

目前限制你的产量的主要因素有哪些?

- 不存在
- 劳动力短缺
- 原材料或设备短缺
- 财务约束
- 其他因素(请具体写出)

这里以匈牙利的调查结果为例(图5和表A1)。在得出与本文有关的结论前,有必要再次看看研究者们到底在问经理人什么问题。为了进行比较,请大家看看上文的第3个假想案例。问题的内容存在本质的差别!这里的研究者不是要求26个国家的经理人预先估计自己的最大产量,假如存在足够多的需求,他们能够生产多少产品。这里的问题不是关于供给能力,而是针对预先估计的产量。研究者也没有问,假定企业有生产500件产品的能力,也有这方面的外部条件,他们是否愿意生产500件产品。或者说,这里没有问他们在不考虑需求的情况下,能实现多大的供给量?

相反,这里的研究者的提问隐含地假设,生产经理在制定产量目标时会考虑各种障碍,并得出一个现实的、可以完成的计划。按照本文的研究思路,这可以在供给形成过程的晚期阶段形成。这里的研究者是在供给形成过程中提问,通常会有哪些因素制约产量,也就是说“假如你的预测做得不错,一般会考虑到哪些障碍因素?”这不是调查中的问题的原话,我将其解释出来,是为了看清问题设计背后隐含的假设。

<sup>①</sup> 调查方法参见 OECD (2003) 和 Nilsson (2001)。

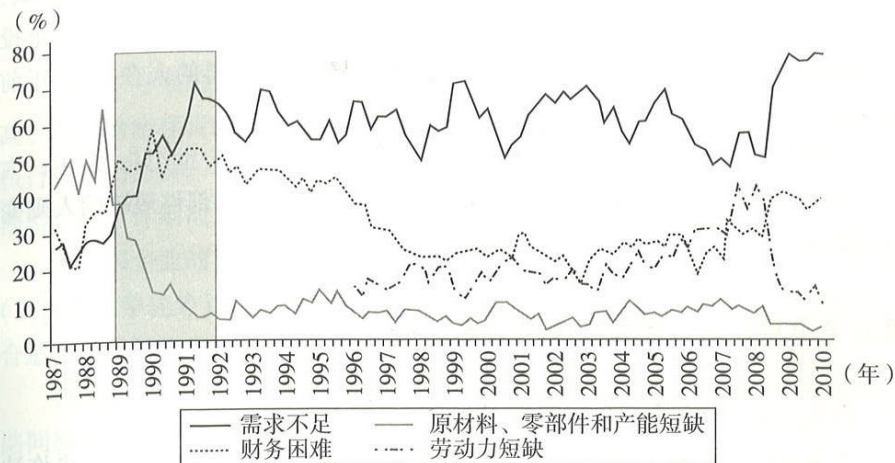


图5 影响匈牙利的制造业产量的因素(1987~2010年)

注:受访者要回答哪些因素影响了产量,可以提到多个因素。为简明起见,本图仅包含少数时间序列的数据。纵轴代表提到某个因素的相对频率。图中的灰色实线代表原材料、零部件和产能的短缺,具体是指,在每个调查期,我们选择五个变量(有关原材料、零部件和产能短缺)中的最高值(最高的相对频率),或者说是至少提到其中一种因素的人的最高比例。图中的灰色区域代表的是体制转轨的时期。数据是每年的第一季度。

资料来源: Direct communication of Kopint-Türki (Institute for Economic and Market Research, Budapest)。该图基于附录表 F.1. 的数据绘制而成。

这个敏感而有趣的问题带来了敏感而有趣的回答,<sup>①</sup>但没有反映出原始的供给能力(会达到实际产能的边界)在微观层面会有何种需求约束。于是,这一调查数据无法让我们直接得到微观层面的过度需求或过度供给的规模,没有也不能为本研究有关过剩经济和短期经济的基本问题提供清楚而直接的答案。

但无论如何,通过谨慎的分析,从这些数据中还是能为我的研究课题找到一些间接的启示。请注意,在匈牙利的数据里,提到投入困难和需求不足问题的人的比例发生了显著变化。图5的灰色区域代表的是经济体制转轨时期。时间趋势线明确显示,到体制转轨之前,社会主义短缺经济中的企业实现连续生

<sup>①</sup> 在这26国进行的调查不是为了测算过度供给或过度需求现象的规模或分布,而是为商业周期波动提供帮助。

产和增长所遇到的障碍主要来自资源方面：劳动力尤其是熟练劳动力的短缺、原材料和零部件的短缺是最常见的问题，有三分之二的受访者至少提到一个投入方面的问题。需求不足的情况并不罕见，但提到此类障碍的人在1989年前不到受访者的三分之一。

在体制转轨后，答案的分布比例基本上颠倒过来。到21世纪头10年，大约三分之二的受访者提到需求方面的障碍，关注原材料和零部件短缺的人则变得非常少。<sup>①</sup>

### 第三个难点：从过剩存货中区分出“必要”的部分

回到讨论主线，我们继续看概念和测算的问题。

在第3章开头的简化案例中，未售出的期末存货被视为过度供给。也许读者并不在意，但我有责任指出，这样的理解存在问题。

在存货和产品更新保持固定节奏的时候，如果某个时期的未售出存货增加，显然存在过度供给。但之前我曾强调，需求并非固定不动，而是在不断变化中。保持一定的存货是为了更好地适应供给和需求的变化。

那么在何种情况下我们可以认为存货过多、出现了持续的过度供给呢？在回答这个问题之前，我再谈谈第1章曾提到的安全水平的概念，这个指标在下面的讨论中用B来表示。例如，人们可以认为，如果90%的客户能在商店找到自己想要的商品，仅有10%必须用其他物品替代或者空手而去（ $B=0.9$ ），那么这家商店的存货水平就是充足的，而没有过多。但是，如果买不到合适商品的客户的比例下降到了9%（ $B=0.91$ ），存货就过多了，存在持续的过度供给。但坦率地说，这样的说法有些主观臆断。为什么B应该等于0.9，而非0.85或0.96呢？安全水平B并不是给定的，而必然是待定的：它的水平在一定程度上要受第2章所介绍的共同驱动和调整机制的影响。<sup>②</sup>

因此，安全水平必然是待定的，越过安全水平之后，存货量就超出了必需，开始出现过多的存货量，可以毫不犹豫地将其视为过度供给。

<sup>①</sup> 有必要指出，熟练劳动力的短缺在资本主义转轨完成中仍然是个问题。

<sup>②</sup> 与此论点一致的是，有些企业会执行明确规范的存货和订货规则，从而影响对客户的安全水平B。

在这一数量分界线待定的情况下，我们仍然可以确信存在某些重要的相互关系。

• 除了习惯的延续和进货的安排外，安全水平随存货的结构合理性而提高。如果给定存货的规模与结构，安全水平随进货的弹性和便利程度及订单的执行速度而提高。<sup>①</sup> 这些联系是显而易见的。人们设计了各种研究商业运营的模式，用于详细地量化分析这些联系。

• 如果企业的存货政策允许实现较高的安全水平，这既会给卖方带来影响（如加强竞争优势），也会给买方带来效应（如提供更多的选择，增加找到最合适商品的机会等）。

以上的讨论仅仅是针对过度供给，因为这是本文的中心议题，但论述的思路同样适用于过度需求的情况。如果所有希望购买、也有财力购买的客户来商店后都空手离开（ $B=0$ ），存货水平显然太小。但如果经验研究显示，实际的需求规模和结构会导致安全水平 $B=0.3$ 或 $0.4$ ，又该得出什么结论？存货水平依然太少吗？我需要重申：存货水平过低的分界线在哪里，这依然是个待定的问题。

回到定义的问题：以上的讨论已足以对“多余”、“过度”这样的特征描述产生质疑。对我来说，我希望在本文中回避是否面临过多存货或过多产能的问题，当然有时未必成功。

在以前的《短缺经济学》（Kornai, 1980）一书中，我采用“宽松”（slack）一词来表示这里所说的“过剩”的含义。后来，我有了更多的时间与讲英语母语的人接触，也阅读了更多英文的日常报道。我现在感到“slack”带有贬义，有松弛的含义，让人感觉至少有部分存货和产能是多余的。于是我一直试图寻找一个更中性的词汇，以避免对判断的误导。我还有意地不做结论，闲置的产能、多余的存货和缓慢的流转是应该被视为浪费，还是属于聪明的储备、对不可预测的需求的特别应对办法。

尽管我坚持认为“过度”的数量分界线是待定的，但并不代表未售出的存货规模以及因此未能满足的客户比例就无关紧要。这些可以观测到的联系通过适当的方法可以进行量化分析。

<sup>①</sup> 日本广泛采用即时生产策略，如果必要的投入品能够及时送达，而不用建立存货，则可以在较低的总体存货水平上维持连续生产。

例如，对存货的构成就可以进行颇有启发价值的分析（如 Chikán, 1984），我们来看看表 3。

表 3 投入存货和产品存货之比，国际比较（1981~1985 年）

国家	制造业投入存货和产品存货之比（1981~1985 年）
社会主义国家	
保加利亚	5.07
捷克斯洛伐克	3.07
匈牙利	6.10
波兰	4.49
前苏联	3.16
资本主义国家	
澳大利亚	1.36
奥地利	1.06
加拿大	0.92
芬兰	1.92
联邦德国	0.71
日本	1.09
挪威	1.10
葡萄牙	1.66
瑞典	0.81
英国	1.02
美国	1.02

资料来源：由 A. Chikán 整理，发表于 *The Socialist System* (kornai, 1992, 第 250 页)。

表 3 的计算基于如下理论：不管是在哪种经济制度下，生产和销售都需要保持连续性以及避免受冲击，因此要保持存货，包括投入的存货和产品的存货。然而，两者之间的比例会反映经济制度的特点：哪方面需要更充足的存货储备？从表 3 可以看到，在社会主义经济制度下存在短缺经济，因为投入品的采购面临很大的不确定性，而普遍的短缺导致销售相对轻松，投入品存货相对于产品存货的比例非常高（Farkas, 1980）。相反，在资本主义经济制度下，

这一比例要低得多。这表明在资本主义制度下，投入品发生短缺的可能性要小得多，而产品存货会更多堆积，因为销售更加困难，生产商希望给买家提供更快服务与更多的选择。

#### 第四个难点：不合理的加总

我们来看由相关的产品或服务组成的局部市场。在标准的微观经济学模型中，习惯将“净过度供给”列入价格变动的解释变量。如果这个变量为正值，价格上涨；如果为负值，价格会下跌（在瓦尔拉斯的经济学理论中有个著名的人物是“拍卖商”，在奥斯卡·兰格的理论中则是价格管理办公室，也是根据拍卖商的规则来调节价格，直到均衡的达成）。

要得到“净过度供给”，只需要把过度需求的值（为正值）和过度供给的值（为负值）简单加起来即可。

初看起来这似乎很符合逻辑，背后却存在重大错误。假如我们是在分析一个机场的运输情况：离开机场的飞机的上座率如何？此时的过度需求是指，有的旅客无法在期望的时间坐上航班。过度供给则是指，有的飞机只搭载了半数的乘客。净过度需求是指，在给定时间内，买不到票的乘客人数减去空座位的数量。然而这个加总值没有丝毫意义，对于那些未能在早上 9 点登上布达佩斯去哥本哈根的航班的乘客来说，告诉他们晚上 7 点的航班还有座位，足以提供安慰吗？他们可能在当天下午有重要的会面。

在理论模型和实证计量中，把过度需求当作净值来处理是常见的做法。<sup>①</sup>但是反思表明，应该严格限制使用“净过度需求”或“净过度供给”这样的指标。

以上讨论的所有测算方面的难点都提醒我们应慎重采用加总的做法。实际上，深入地思考表明，应该完全禁止这样的加总。让我们重新思考一下上面讨论的几个问题，纯粹从加总的视角来看：

(1) 不可能把期末存货与闲置产能加总起来，因为它们代表着不同性质的供给可能性。在这两种过剩中，闲置的产能更为重要，因为存货的规模更容

<sup>①</sup> 加总的问题在社会主义国家宏观经济分析所利用的非均衡模型中尤其突出，参见 Portes 和 Winter (1980)；Portes 等人 (1987)。有关讨论参见 Davis 和 Charemza (1980)；Brabant (1990)。

易控制，而产能更为持久，也更难缩减。

(2) 供给和需求的形成都是动态的过程。每个卖家或买家的销售倾向或购买倾向在过程中都可能发生变化或成为事实。在某个时间点进行横截面测算的时候，把尚未完成的不同性质的购买倾向或销售倾向加总起来是不对的。

(3) 最后一个或许也最严重的问题在于，利用“净值”来对过度需求和过度供给进行抵消，是完全错误的做法。

以上观察当然也适用于更低或更高的分析层级。这一论述思路直接表明，在解释宏观层面的数据时需要最大的谨慎。对于某个特定时点的某个特定经济体，不可能得到宏观层面的供给和需求数据。

必须承认，我不太清楚应该在哪里阐述自己对实际 GDP 和潜在 GDP 之“差额”的思考。这个指标在宏观经济分析中被广泛采用，同样是反映过度供给的指标，类似于某个生产企业或某个产业的过剩产能。看起来，上文提到我对更低层面的加总指标的担忧和反对，对这一宏观指标同样适用。这个问题显然还有待进一步澄清。对我而言，我现在既不敢接受也不敢否定这个被广泛采用和重视的指标。

我还可以列举更多统计和观测上的困难和异常状况，但以上缺陷已足以解释我为什么愿意用“短缺经济/过剩经济”这一组概念来描述特定国家的普遍市场状态，而不是采用“过度需求经济/过度供给经济”这一组许多人认为更容易理解的概念。

## 测算和概念工具的实用建议

根据上文的论述，我已不指望测算出在给定市场和给定价格水平上，总供给和总需求的水平是多少，不指望能量化估计过度供给或过度需求的规模。我不会尝试用一个关键指标来总结某个市场的普遍状态。

但这并不代表彻底放弃测算任务。虽然不能用汇总的关键指标，但仍有可能利用几个局部指标来反映过剩经济和短缺经济的典型特征，以及它们在给定时间的状态及其变化。

这里先列出此类指标的几个例子。

- 生产商和服务商的产能及其利用率。

- 存货周转率，存货的构成比例的相对变化以及相对于销售额的变化。
- 让生产商回答，生产面临的主要障碍是哪些？
- 排队和等待的时间，排队者的购买意愿与实际销售量的对比。

以上指标的例子会在本文的好几个部分出现，表明它们是在实践中可行的测算办法。

也可能找到其他指标来反映过剩和短缺现象的某些特征。我们的富有创造力的经济学家、市场研究人员和统计学家们可以设计出更多的观察与测算方法。但不幸的是，往往不能从实际的观察和测算中总结出理论，这一理想过程在实践中经常是以相反的次序完成：把某个基于假设的理论作为出发点，然后对现象进行观察和测算，那时才能进行统计分析。

上文提到的所有测量方法会得出不止两种不连续状态：过度供给与过度需求。它们会报告某个现象的频率或程度：闲置产能已达到 10% 或 30%，存货可以满足 3 个月还是 23 个月，排队等待的时间是 3 个月还是 3 年，这些并非无关紧要。

根据这些信息和类似指标，我们有可能对过剩现象和短缺现象的分布及其变化规律进行分析（还可以根据它们的频率或程度进行细致的分类），发现分布的随机特征。通过对这些指标的仔细研究，可以对某个给定市场的一般状态得出有说服力的结论。

## 编制合成指标或“综合指数”

我们还可以尝试，对涉及某个国家整体经济的过剩现象或短缺现象编制出“综合指数”。这样的综合指标可以用于几方面的目的，其中最常见的例子包括：

- “自由指数”是为了反映某个国家在某个年份的企业和个人权利的自由状态（Freedom House, 2010）。
- “腐败指数”反映的是某个国家在某个年份的腐败事件发生的范围和严重程度（Transparency International, 2010）。
- “商业环境指数”反映的是商业领域的决策者的“心情”，对于自己预期的乐观或悲观程度（参见 Clavel 和 Minodier, 2009；Erkel-Rousse 和 Minodier, 2009）。



这些测算的原始目的是某种现象的总量或频率，但难以实现直接测量，只能考虑测量一系列局部现象。于是人们设计出这样的函数：解释变量是反映局部现象的局部指数，被解释变量则是“综合指数”。<sup>①</sup>

这样的“综合指数”可以通过一定的程序从局部的构成指数中求得。最简单当然也最粗略的做法是计算出局部指标的平均值，作为综合指数。有时会采取因素分析的方法，权重最大和解释力最强的因素被当作综合指数。

本文的目的不是设计能够对系统性的短缺或过剩进行测算的“综合指数”。我甚至都不愿意承认，通过对那些现象的简明思考肯定能得到这样的指数。这一工作需要反映市场状态的局部指标的变化进行深入分析。我只是想强调，简单的加减并不适合计算那些系统性的复杂现象，它们在统计上很难追踪，有必要重新借鉴人们在处理不能由关键指标反映的其他综合现象时的经验。

在本章结尾，回顾对澄清概念和测算方法的讨论，我要再次强调概念和测算工具保持“中性”的重要性，它们既可以应用于过剩现象占据优势的市场，也适用于短缺现象占据优势的市场。在观察和测算市场的实际状态时，保持中性也是应该的，而不要受人们自己对某个经济制度的价值判断或爱憎情绪的影响。

#### 4. 劳动力市场：不断产生过剩的机制

##### 概念定义和测算方法

在对劳动力市场进行深入分析前，需要对概念和有关测算方法进行梳理。我希望沿用产品和服务市场采用的办法，因此，必须弄清楚本文的概念工具与统计指标之间存在何种持续联系。<sup>②</sup>

<sup>①</sup> 我在很早以前的研究中首次提到这点（*Growth, Shortage and Efficiency*, Kornai, 1982）。后来，我和几位同事开始了一项重要研究，设计短缺的局部指标，最终测算了反映短缺的综合指数。体制转轨到来后，该项目随之结束，参与者转向了其他研究。匈牙利短缺经济的终结使这个研究失去了迫切性。直到现在思考过剩现象的综合测算时，我才想到应重新考虑。

<sup>②</sup> 国际上通用的定义参见：International Labor Organization (2010) 版。

在出现短缺现象的时候，情况相对简单。统计上报告的岗位空缺数就是劳动力短缺很好的反映，另外还报告有登记失业人数，显然是劳动力过剩的基础数据之一。然而，在劳动力市场过剩的情况下，量化分析需要做更多工作。

首先，根据本文的宗旨，我们要搞清楚需要测算哪些指标，并将其与现有的数据进行对比。我们的研究是针对某个经济体，针对给定时间内全国范围的情况。我们假定这个国家的总人口数为  $Q$ ，可以划分为四个群体：

(1) 有些居民在任何情况下都不会就业。例如，儿童不能参加工作，只有 14 岁以上的人在统计上被视为有劳动能力的人。但经济史也表明，用法律来禁止童工是近代才出现的事情，许多落后国家今天仍然有大量儿童从事工作。有的居民由于身体或精神上的限制也不能参加工作。也需要注意，不能把官方承认的身体或精神障碍的所有情况都自动视为没有劳动能力的人，因为其中很多人实际上可以从事工作，只是缺乏必要的社会和经济条件。这其中包括那些由于疾病及身体或精神障碍而被工作排斥的老人。然而，并非所有的老人都不能工作，哪怕超过官方的劳动力统计所划定的最高工作年龄的人。例如，本文的作者写作时尽管已年满 83 岁，仍然感觉自己完全有劳动能力。以上每个没有劳动能力的群体后面的补充内容表明，这些标准在一定程度上都存在疑问。我们把不能参加工作的居民人数用“ $N$ ”来表示。

(2) 第二个群体是指有劳动能力，但由于某些原因没有进入劳动力市场去寻求就业机会的人。我们称这个群体的人为“不积极劳动力”，用“ $M$ ”表示。有各种原因影响一个有劳动能力的人去寻找工作：

- 可以从其他途径获得生活来源（私人财富、其他家庭成员的收入或者政府救济等）。
- 可以依靠养老金生活的退休人员。
- 由于习惯而不愿意参加工作，对女性而言尤其突出。
- 女性就业还受到其他因素的影响。特别是需要在她们的个工作日内解决看护问题，如托儿所、幼儿园、放学后的看护中心、老年人看护机构等。
- 在很长时间内寻求就业未获得成功，于是停止找工作。这些人被称为“沮丧的劳动力”（discouraged workers），在匈牙利的劳动统计中视为“消极失业者”（passive unemployed）。

这些打击因素在一定程度上重叠，很难说那些不去找工作的人在多大程度上是出于自愿还是因为环境的束缚，这个问题处于经济学、社会学和社会心

理学的边缘地带。束缚可能是来自社会习俗（女性就应该照顾家庭的观念），也可能是缺乏工作所需要的环境条件（没有托儿所或费用太高）。那么，一个找工作的人最后放弃希望，“自愿”退出劳动力市场的边界点到底在哪里？

我得承认，以上两个群体，即没有劳动能力和有劳动能力但不积极参加工作的群体的分界线不见得是绝对的，但他们之间还是存在明显的重大区别。

(3) 官方统计中记录的“登记失业者”，用“U”来表示。政府采用严格的标准来区分积极找工作的失业者和放弃找工作“消极失业者”，例如在接受询问前的四个星期里是否曾积极寻找工作等。很自然，这里存在主观臆断的危险，为什么非得是四个星期，而不是三个或五个星期呢？当然，这样的主观判断对任何统计记录过程都是难以避免的。

(4) 最后，官方统计中记录的在某个时点正在就业的群体，用“E”来表示。

在官方统计中，就业者和积极找工作的失业者之和被称为人口中的“经济活跃群体”，在英文的研究文献中称其为“劳动力”，其数量用“A”表示， $A = E + U$ 。人口中其他部分被称为“非劳动力”，用“B”表示， $B = Q - A$ 。

从本文的宗旨出发，过剩劳动力（用“T”表示）不但应该包括官方登记的失业者，也涉及人口中有劳动能力但不活跃的部分， $T = M + U$ 。借用马克思主义的术语来说，他们是“劳动力后备军”，在市场需要的时候可以被重新招募。在商业周期的上升阶段，市场会吸纳一部分过剩劳动力，不但会导致失业人数减少，还会让一些已经退出市场的人重新变得活跃起来。战争期间这个现象更加明显，所谓的后备军会被真正地动员起来，政府将强制人们参与生产性的活动。

官方统计在区分经济活跃人群与不活跃人群时有不同的划分标准，表4对这些概念上的区别做了更清楚的总结。

本文的概念框架在一个关键点上与官方的统计定义有所不同：政府并没有经常报告有劳动能力但不活跃的人群数量。在某些研究文献里，出现过“能参加工作的年龄段”的人群，通常是指14~74岁的人。但这样的分类与我的研究目的也不相同，因为在这个年龄段中的某些人绝对不会参加任何工作，而在74岁以上的人群中也有某些人有劳动能力，但不处于活跃状态。

在下面的讨论中，我假定用统计上可以观察的指标“b”代表经济上不活跃的人群所占的比例，并把它作为指标“m”的替代指标。“m”是指有劳动能力但经济上不活跃的人群所占的比例，符合本文的研究目的，但在统计上缺乏数据。利用这个指标，我将在下面对一个特定国家的时期变化进行比较分析，寻找广泛适用的制度性特征。我的假设主要是基于对问题的逻辑思考，我推测不能参加工作的人数是相对固定的，主要取决于人口结构和健康指标。另一方面，在有劳动能力的人中间，是否积极参与经济活动的选择主要取决于社会和经济因素。选择的结果非常清楚地反映在我建议的指标“M”中（目前没有做直接统计），但官方统计等于是在这个变量上面增加了一个相对固定的数值（不能参加工作的人口数量“N”），反映的是其总数“B”， $B = N + M$ 。

表4 本文所使用的概念与劳动力市场统计数据使用的概念之间的联系

本文所使用的概念	劳动力市场统计数据使用的概念	
	能参加工作的人口 (N)	
过剩劳动力 ( $T = M + U$ )	不能参加工作人口 (M)	经济上不活跃的人口 (B)
	失业人口 (U)	
		经济上活跃的人口 ( $A = U + E$ )
短缺劳动力	就业人口 (E)	
	空缺岗位 (V)	空缺岗位 (V)
总人口 ( $Q = N + M + U + E$ )		总人口 (Q) ( $Q = B + A$ )
相对指标 (%)		相对指标 (%)
过剩比例 ( $t = T/Q$ )		经济上不活跃的人口比例 ( $b = B/Q$ )
有工作能力但经济上不活跃的人口比例 ( $m = M/Q$ )		经济上活跃的人口比例 ( $a = A/Q$ )

续表

本文所使用的概念	劳动力市场统计数据使用的概念
失业率 ( $u = U/Q$ )	失业率 ( $u = U/Q$ ) 或 ( $u' = U/Q'$ )
短缺比例 ( $v = V/Q$ )	空缺岗位与失业人口之比 $v = V/Q$ , 或 $v' = V/W$ , $W$ 表示可供就业的岗位 总数

注：官方的劳动力市场统计通常会报告中用括号标注的指标，如  $u'$  和  $v'$  等，不过  $u$  和  $v$  的数据也很容易推算出来。在测算劳动参与率时，官方的劳动力市场统计并没有将人口总数  $Q$  作为分母，而是采用的  $Q'$ ，代表 14~64 岁的人群中有劳动能力的人。

未来的一个研究任务将是在实证研究中区分没有劳动能力的人群与有劳动能力但不活跃的人群，这些数据是可以观测的，并没有什么测算上的困难。

### 体制转轨给劳动力市场的冲击

在本文开头，我提到波兰的一个场景，短缺经济时期的人们在食品商店门口排起长队，货架上空空荡荡。在今天的波兰和其他后社会主义国家，货架上的商品已经琳琅满目，顾客们也对此习以为常。

劳动力市场出现了同样的变化。经济上较为发达的中东欧国家和前苏联在 1989 年以前遇到过严重的劳动力短缺问题，体制转轨则给劳动力市场带来巨大冲击，大量工作岗位消失，失业长期持续。几年时间里，经济活跃的人数显著减少。在产品市场上，人们很容易接受过剩经济的出现，但在劳动力市场却很难，这一转变经常令人难以忍受。

对于我这样的东欧经济学家而言，亲身经历这样的变革迫使自己重新思考劳动力市场的状态与经济制度之间存在的关系。西方国家的社会科学研究者对这一问题的冷漠态度令人遗憾，我甚至认为，没有哪位关注劳动力市场的西方学者在探寻失业原因的时候曾试图对资本主义与社会主义的历史进行比较研

究。西方的学术会议讨论者认为，即便在所谓充分就业的情况下，也会存在失业，这是不需要证明的。但直到今天，每当我读到经常被引用的“自然失业率”的说法时，仍然不禁会感到烦心甚至愤怒。<sup>①</sup> 自然失业率？由森林、动物、岩石和地震这些成分组成的大自然曾经宣布过，失业是应该存在的吗？数十年来，我一直在尖锐地批评社会主义经济制度，但无论是社会主义的朋友还是对手都应该注意到，在这样的经济制度下存在长期的劳动力短缺，而不是长期失业和大量的劳动力过剩。

接下来我想用一些数据来描述后社会主义国家的劳动力市场经历的剧变。不幸的是，没有时间序列数据可以反映体制转轨前的劳动力市场的状况，与之后的 20 多年用同样的测算指标进行对比。不过，从这些图示和表格中依然能明显看到变化的剧烈程度（图 6 来自 Kornai, 1992, 第 209 页）。

图 7 反映的是图 6 中的指标在 2009 年的水平。

请注意图 6 中的劳动参与率，与类似发展水平的资本主义国家相比，所有社会主义国家（图中用虚线包围的部分）的水平都更高。

对比图 6 和图 7 则可以看到体制转轨带来的显著变化：用虚线包围的小集团被打破了，前社会主义集团的国家的劳动参与率指标不再显得特别突出（它们目前的状况用小圆圈代表）。可以看到，这些国家与其他国家已混合在一起，基本上与经济发展阶段类似的资本主义国家接近，还有个别国家的劳动参与率偏低。

不幸的是，以前没有做过覆盖多个国家的综合调查，以测算当时的劳动力短缺程度。至于其他数据，这里找到一个基于波兰的调查的图表（来自 Kornai, 1992, 第 242 页）。图 8 的时间序列终止于 1988 年，该曲线显示了空缺岗位相对于失业人数的比例。以对数距离表示的纵轴表明，在 1988 年，相对于每位找工作的人，有 86 个以上的空缺岗位。

表 5 显示的是经济不活跃人口、失业人口以及空缺岗位的时间序列，包括若干前社会主义国家及其当前的情况。

<sup>①</sup> 请注意，这里的议论并不针对“自然失业率”理论的内容。引起我烦恼的并不是研究本身，而是这个术语。

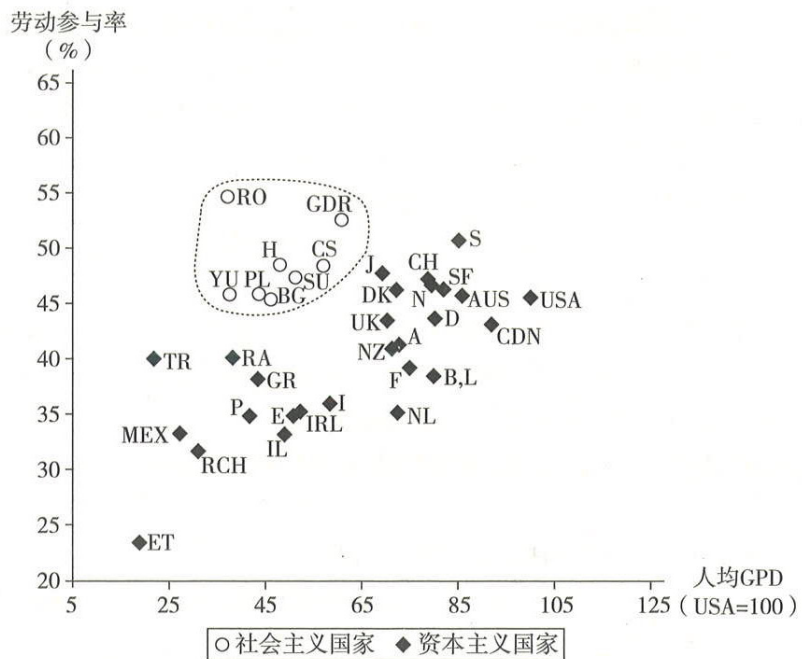


图6 劳动参与率与经济发展水平（1980年）

注：图中包含的国家包括（按人均GDP为序），埃及（ET）、土耳其（TR）、墨西哥（MEX）、智利（RCH）、罗马尼亚（RO）、南斯拉夫（YU）、阿根廷（RA）、葡萄牙（P）、希腊（GR）、波兰（PL）、保加利亚（BG）、匈牙利（H）、以色列（IC）、苏联（SU）、西班牙（E）、爱尔兰（IRL）、捷克斯洛伐克（CS）、意大利（I）、民主德国（GDR）、日本（J）、英国（UK）、新西兰（N）、丹麦（DK）、奥地利（A）、荷兰（NL）、法国（F）、瑞士（CH）、挪威（W）、比利时（B）、卢森堡（L）、联邦德国（D）、芬兰（SF）、瑞典（S）、澳大利亚（AUS）、加拿大（CDN）和美国（USA）。社会主义国家用圆圈表示，资本主义国家用菱形表示。纵轴表示的劳动参与率的计算方法是：经济活跃人口与14~84岁的有劳动能力的人口之比。

资料来源：Kornai（1993，第209页）。科洛（János Köllö）负责收集数据并绘制图示，人均GDP的数据来自一篇论文（Éva Ehrlich，1985，第100页）。资本主义国家的人均GDP和劳动力市场统计数据来自ONU编辑的一本年鉴，社会主义国家的数据来自克姆孔（Comecon）编辑的一本年鉴。

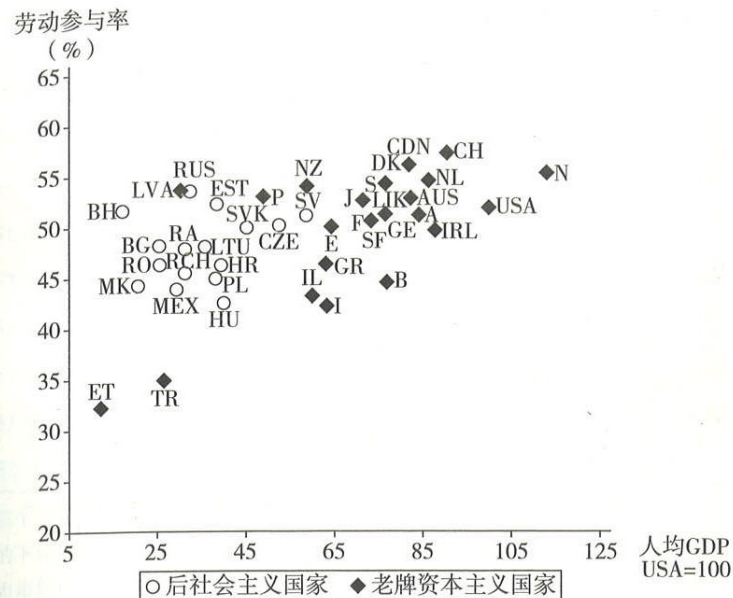


图7 劳动参与率与经济发展水平（2009年）

注：图中包含的国家包括（按人均GDP为序），埃及（ET）、波黑（BH）、马其顿（MK）、保加利亚（BG）、罗马尼亚（RD）、土耳其（TR）、墨西哥（MEX）、拉脱维亚（LVA）、智利（RCH）、阿根廷（RA）、俄罗斯（RUS）、立陶宛（LTU）、爱沙尼亚（EST）、克罗地亚（HR）、波兰（PL）、匈牙利（H）、斯洛伐克（SVK）、葡萄牙（P）、捷克（CZE）、斯洛文尼亚（VS）、新西兰（NZ）、以色列（IL）、希腊（GR）、意大利（I）、西班牙（E）、法国（F）、日本（J）、芬兰（SF）、德国（GE）、英国（UK）、瑞典（S）、丹麦（DK）、比利时（B）、澳大利亚（AUS）、加拿大（CDN）、奥地利（A）、荷兰（NL）、爱尔兰（IRL）、瑞士（S）、美国（USA）和挪威（N）。后社会主义国家用圆圈表示，老牌资本主义国家用菱形表示。纵轴表示的劳动参与率的计算方法是：经济活跃人口与14~84岁的有劳动能力的人口之比。卢森堡的情况比较特殊，没有纳入本图。

资料来源：人均GDP来自World Bank（2010）；劳动参与率的计算数据来自ILO（2010）。

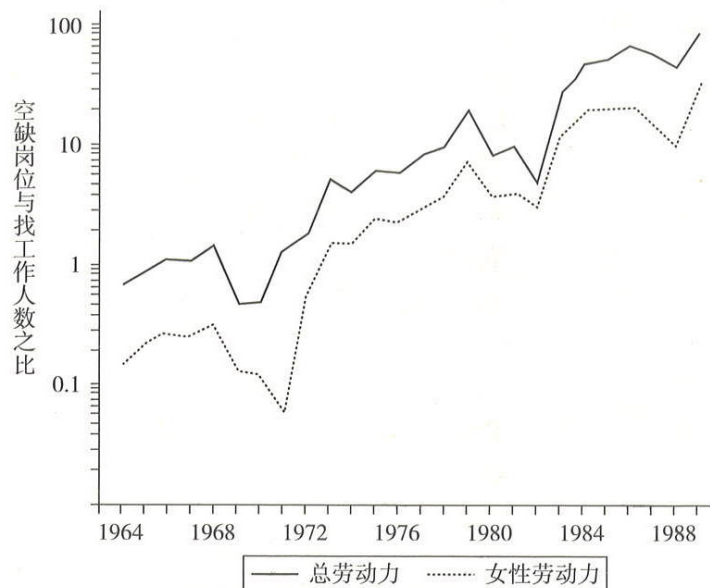


图8 相对于每位找工作的人的空缺岗位数 (1964~1988年)

资料来源: 该图来自 Kornai (1993, 第215页)。由 János Köllönd 基于 Fallbuchl (1982, 第33页) 和 R. Holzmann (1990, 第6页) 的数据绘制而成。

表5 东欧国家的经济不活跃人口、失业人口以及空缺岗位 (1989~2008年)

年份	捷克			爱沙尼亚		波兰			拉脱维亚		
	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>v</i>
1989	53.34			47.10	0.30	54.05		0.12	46.49		
1990	53.05		0.40	47.89	0.34	54.45		0.12	47.39		
1991	52.18		0.75	48.73	0.78	54.77		0.08	48.35		
1995	50.77	2.02	0.96	52.50	4.73	56.08	5.90	0.08	52.72		
2000	50.47	4.45	0.57	53.30	6.57	55.62	7.25	0.03	55.04	6.62	
2005	49.98	4.02	0.83	52.01	3.87	54.75	7.97	0.12	52.18	4.32	0.53
2006	49.96	3.63	1.21	50.58	3.01	55.18	6.14	0.17	50.88	3.50	0.84
2007	50.22	2.69	1.36	50.25	2.38	55.06	4.25	0.17	49.72	3.13	0.89
2008	50.40	2.23	4.46	49.47	2.86	54.35	3.18	0.10	38.53	4.01	0.47

	立陶宛		匈牙利		斯洛伐克		斯洛文尼亚			
1989	49.30		56.14		51.69		59.02			
1990	49.65		56.38		51.50		58.88			
1991	50.01		56.83		0.21	50.71		58.56		
1995	51.63		59.96	4.03	0.37	53.87	6.05	52.63	3.46	
2000	52.98	7.68	59.49	2.62	0.42	52.13	9.03	52.68	3.07	
2005	53.53	3.89	0.23	57.91	3.00	0.41	50.63	7.94	50.40	2.85
2006	53.88	2.64	0.58	57.46	3.15	0.30	50.67	6.55	50.27	2.94
2007	53.42	2.06	0.58	57.46	3.15	0.30	50.67	6.55	50.27	2.94
2008	52.92	2.84	0.67	57.75	3.29	0.21	50.14	4.4	49.8	2.26

注: 三个指标都是根据劳动力市场指标和总人口计算得出的。根据表4的定义, 表中的三列包含每个国家的下列指标: 经济不活跃人口比例,  $b = B/Q$ ; 失业率,  $u = U/Q$ ; 空缺岗位率,  $v = V/Q$ 。这三个劳动力市场指标都根据总人口数进行了调整, 以进行更好的比较。爱沙尼亚、斯洛伐克和斯洛文尼亚没有官方的空缺岗位统计。表中只显示了整个时间序列数据的部分观测值。

资料来源: 经济不活跃人口数、失业人口数和总人口数来自 ILO (2010); 空缺岗位数来自 OECD (2010); 拉脱维亚和立陶宛关于空缺岗位的时间序列数据来自 Latvijas Statistika (2010) 和 Statistikos Departamentas (2010)。

该表清晰地反映了下面的变化过程:

• 体制转轨之后出现了劳动参与率的剧烈下降, 具有劳动能力但退出劳动力市场的人数增加, 可能出于几个方面的原因。其一是, 人们发现难以找到合适的岗位, 失望情绪蔓延。

• 该地区的大多数国家在体制转轨前没有持续进行失业统计, 只是零星地做过几次。<sup>①</sup> 不过, 这个指标在 1989~1990 年之后快速提高, 达到与西方国家接近的水平。

• 我已介绍过所能找到的反映体制转轨前劳动力短缺状况的一个指标。即使在体制转轨后最严峻的年份里, 空缺岗位仍然会出现, 但表5中反映出来的空缺岗位数已变得很少: 劳动力市场的短缺现象不再会密集出现了。

<sup>①</sup> 这是指工厂之外的失业, 而非“工厂大门内的失业”, 后者指有的员工拥有正式的雇佣合同, 却并不认真工作的情况。

## 凯恩斯式失业

为什么经济较为发达的社会主义国家会出现劳动力短缺，而且这种现象会长期维持？这不是因为执政当局希望执行“完全就业”的政策。在这个现象背后不存在统治集团的特殊决策，或者是作为中央计委瞄准的目标。劳动力的短缺源自这个体制内在的特征和倾向。

我这里用非常简化的形式描述一下自己以前的研究所做的分析（特别是《社会主义体制》）。这些社会主义国家的企业都是国有企业，其领导人深受扩张冲动（凯恩斯的动物精神的一种表现）的影响，对投资有着无法满足的胃口：企业的经理人希望完成尽可能多的投资，而毫不计较成本。他们对投资的过度偏爱是由于缺乏硬性的预算约束。尽管负责投资资源配置的官僚体系会形成某种形式的制约，还是会发生过度投资，投资失败造成的损失则由政府来承担。

这样的投资饥渴已几乎足以解释：增长和上级机构促成的增长冲动迟早会吞噬全部劳动力储备。社会主义国家早期的投资增长可以依赖从农业向工业的大规模劳动力转移，还有让家庭妇女更多走向就业，但这些储备劳动力迟早会被全部吸干，增长最终会遭遇劳动力瓶颈。

在经典的社会主义经济制度下，政府对价格和工资实行严格控制。劳动力市场的紧张情况不会像市场经济里那样造成显著的通货膨胀效应。虽然价格和工资也会上涨，但由于软预算约束的影响，企业对成本的敏感度远远落后于资本主义制度下的同行。

同时，创新的推进则远远落后于资本主义制度。劳动生产率陷入停滞，或者仅能出现间歇性的增长。扩张冲动要求增加劳动力，可是劳动力的经济活跃度存在社会可以忍受的一个极限。于是，劳动力短缺成为了大规模的长期现象。<sup>①</sup>

这与凯恩斯所描述的失业机制形成了镜像或反例（但稍后我会谈到，准

<sup>①</sup> 事实上，“长期”这个说法有疑问，因为劳动力短缺情况下的经济增长并未维持很长时间，到可以使用“长期”一词的时候，原有的体制已走向崩溃。这其实也是旧体制崩溃的原因之一，劳动力储备的枯竭严重影响了社会主义制度下的经济增长速度，又不能实现创新和生产率的快速提高，自然就导致停滞。

确地说并不是反例）。在社会主义国家，动物精神及其带来的宏观需求膨胀是由如下因素导致的：（1）国有制；（2）政府支付成本，预算约束软化；（3）对价格和工资进行管制。如果这三个关键因素被历史进程改变，如下因素占据主导地位：（1）私有制；（2）政府拒绝扶持亏损项目，强化预算约束；（3）价格和工资市场化，那么生产商虽然还是希望生产更多产品，也希望有更多的劳动力来进行生产，但他们会遇到需求和财务上的约束，这些约束将限制他们无限度地把生产扩大到劳动力供给的极限。<sup>①</sup> 即使剩余劳动力增加，也会在一一段时间的迟滞后才能反映到雇员的工资水平上，工资依然具有粘性（参见 Blanchard 和 Gali, 2007）。

在使用“镜像”这一比喻的时候，我同时注意到它并不十分确切，至少从学说史的角度看是如此。凯恩斯的理论把失业作为周期现象：周期性的经济波动导致需求不足，从而引起就业率下滑。而我的分析是对长期现象的对比：经济发展进入成熟阶段的社会主义经济表现为长期劳动力短缺，资本主义市场经济则存在长期劳动力过剩。即使在经济上升期，大多数西方国家依然存在劳动力过剩：积极寻找工作但未能得到岗位的劳动力，以及有劳动能力却没有登记为失业的人，社会和经济环境的变化有可能让他们重新回到就业中。凯恩斯的理论有助于理解这一长期现象背后的因果机制，而我的结论强调劳动力过剩存在长期性、持续性的特点，与他的讨论对象并不一致。

还有其他机制也对资本主义经济制度下劳动力过剩现象的产生和延续发挥了作用。有些机制同时发挥作用，有些存在重叠、相互影响。<sup>②</sup> 这些解释机制将在稍后进行介绍。恐怕几乎所有人都会同意，其中一个劳动力市场上的作用机制正是凯恩斯式失业（Keynesian unemployment）产生的原因，它表现为一种不变的作用方向：失业在逐渐增加，经济活跃人群的数量在减少。当然，这种作用的加速或减速会受到其他因素（如政府的经济政策）的影响。劳动力过剩增加带来的效应显然被需求和预算约束的放松所缓解了。

<sup>①</sup> 参见非均衡模型：Benassy（1982）和 Malinvaud（1977）。

<sup>②</sup> 有关就业、工资和劳动力市场的研究成果很多。本文并不试图对其进行总结或者挑选，文中提到的研究都是与本文的主题直接相关的。

## 结构性失业

所谓结构性失业是深藏于资本主义制度内的一种机制导致的。<sup>①</sup> 资本主义经济的运动和创新过程（第2章详细讨论的创造性破坏）在不断产生新的就业岗位，同时也在不断消灭旧的岗位。但这两种运动过程并不完全同步。在某个地方失去工作的人未必能找到在其他地方新创造的工作，他们也未必符合新岗位的要求。

产品中应用的新技术要求掌握某些特殊技能，员工们目前掌握的技能在许多方面可能不符合要求，于是，有些技能过时的员工将成为过剩劳动力。适应和掌握新技能需要花费时间，在此期间你可能找不到工作。还有，有些雇主也不能适应新技术，同样有可能退出劳动力队伍。

经济的运动和创造性破坏持续施压，迫使老企业退出市场，这也会造成就业岗位的减少。虽然新加入的企业在不断涌现，它们需要新的劳动力，但这样的需求不见得能与新的剩余劳动力的供给相匹配。

劳动生产率的提高会带来多方面的影响。常见的一个是用资本密集度更高的设备来替代劳动力。随着生产率的提高和经济的发展，生产也会大幅扩张，创造新的就业岗位。不过，这两方面的变化可能出现几种不协调，扩张的出现可能晚于过剩的产生，二者在地理分布上也可能出现差异。

一个明显的例子是每个国家内部的农业转型。各国的村庄里都会出现大量的剩余劳动力，而城市化水平较高的地区却不能足够快地吸收他们。

劳动力需求的重组不但出现在国家内部，也发生在各国之间。随着某些落后国家发展的提速（如目前的中国和印度），这些地区的廉价出口商品对发达国家的许多厂商形成了挤压。中国和印度的农村劳动力进入城市，进入工厂，德国和比利时的员工则在失去工作。这是全球化的必然产物。

于是，我们得出了本文最重要的结论之一：资本主义经济越有活力，这个制度的优点越是突出，结构性失业问题就越严重，制度上的不利倾向也越明显。

活力与创新是资本主义经济的基本特征，但必须补充说，技术进步的方向在事前往往无法预知。1930年代，当西方国家忙于建设钢铁强国的时候，谁会预见到半个世纪之后将出现一个以信息和通讯技术为基础的世界，需要何种

<sup>①</sup> 对于“结构性”的理解并没有一致看法。有人还把“结构性”和“摩擦性”失业视为几乎相同的概念。它们没有严格的分界线。在本文中，我把“结构性”失业理解为与熊彼特的创造性破坏有关的现象，涉及产品、技术、部门、地区和国家之间的资源的重新配置。

不同的专业技能？谁又曾预见我们今天看到的生产在全世界范围内的快速重组？资本主义的辉煌、成就与快速扩张，同时又伴随着千百万人失去自己的工作岗位，仍然在岗的人也面临着更多的不确定性。

## 不匹配调整、摩擦性失业与需求

劳动力过剩和失业的出现还来自劳动力的供给与需求之间的摩擦。寻找岗位的劳动力的技能与空缺岗位需要的技能往往存在差异。在有些情况下，这几乎是不可能克服的。而在其他一些情况下，劳动力的卖方和买方能实现匹配，却没有机会取得联系，这方面的问题与劳动力市场的信息流情况密切相关。失业者在寻找合适的岗位，雇主在寻找合适的员工，这是个费时费力的过程，在此期间，寻找工作的人没有从事职业活动（有关摩擦性失业和寻找理论的研究，可参见 Kornai, 1970; Phelps 等人, 1970; Diamond, 1982; Pissarides, 2000, 第5章和第6章）。

各种社会经济制度下都会出现此类不匹配现象，无论劳动力市场是表现为短缺还是过剩。适应性摩擦是一种注定要产生摩擦性失业的运动机制，从某种意义上讲它并不是某种制度特有的现象。然而，这样的机制仍然会受到外部环境，即市场整体状况的影响，下文还将详述。

上节分析的结构性失业和由于信息缺乏与搜索时间造成的摩擦性失业之间有着密切联系，在一定程度上还存在交叉。如果市场双方都拥有完善的信息，包括对未来发展的准确预测，结构性变革带来的不匹配现象就会减少。反之亦然：如果结构性变化（技术、产业构成和地理分布等）停止，寻找工作与现实供需匹配也将变得更容易。

在中央高度控制的官僚主义的社会主义经济中，技术革新速度很慢，以上两方面的问题都会出现，但主要问题还是由于技术发展的停滞造成的。而对资本主义经济来说，即使市场双方的信息都非常通畅，活跃的技术进步本身还是会造成大量失业。因此，我将其视为资本主义的本质特点，而信息问题所导致的摩擦性失业则不同。<sup>①</sup>

<sup>①</sup> 由于信息不畅导致的摩擦在两种制度下都存在，但在资本主义制度下，人们对于失误所造成的损失会有更直接的感受。效率较低的社会主义企业在亏损状态下继续运转（软预算约束），雇员受劳动力市场的摩擦影响较小，因为工资差距普遍较小，在他们愿意的时候，也更容易换工作。

上文讨论的存在部分重叠的两个现象——持续的结构重组和匹配所导致的摩擦——使劳动力市场上同时出现了短缺和过剩现象。两种现象的同时发生在第2章和第3章也曾提到过几次，有关产品、服务和产能的配置。这里我们看到劳动力市场上的情况也是如此。我们可以从表6的统计数据中得到更多启发，其中三个国家的时间序列数据还以图示的形式表现为图9。

表6 经济不活跃人口的比例、失业率和空缺岗位率的国际比较 (1989~2008年)

年份	日本			西班牙			瑞典			英国			美国		
	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>v</i>
1980	54.07	0.98	0.32	62.87	4.08	0.03	47.97	1.16	0.36	53.88	0.00	0.17	51.96	3.33	
1982	53.60	1.12	0.31	62.70	5.69	0.05	47.56	1.85	0.25	53.62	0.00	0.26	51.47	4.57	
1984	53.09	1.32	0.33	62.37	7.29	0.10	47.21	1.85	0.44	52.90	5.15	0.29	50.96	3.58	
1986	52.86	1.37	0.36	62.95	7.72	0.12	46.77	1.52	0.55	51.56	5.60	0.41	50.23	3.39	
1988	52.34	1.27	0.51	61.04	7.51	0.15	46.46	1.01	0.59	50.80	4.46	0.39	49.82	2.69	
1990	51.19	1.09	0.52	60.28	6.46	0.11	46.01	0.99	0.22	50.14	3.51	0.21	49.72	2.76	
1992	50.27	1.14	0.38	59.83	7.39	0.05	48.03	3.00	0.10	50.99	4.85	0.22	49.94	3.68	
1994	50.20	1.54	0.38	58.86	9.88	0.07	49.59	4.82	0.17	51.62	4.74	0.32	49.99	2.99	
1996	50.04	1.80	0.44	58.22	9.26	0.13	49.77	5.00	0.18	51.72	4.03	0.49	49.73	2.64	
1998	50.03	2.19	0.39	57.13	8.00	0.36	50.67	4.15	0.29	51.60	3.06	0.60	49.32	2.21	
2000	50.62	2.55	0.47	55.42	6.18	0.29	50.62	2.95	0.31	51.05	2.73	0.72	49.22	1.98	
2002	51.17	2.83	0.53	54.78	5.05	0.36	49.54	2.67	0.28	50.95	2.48	1.02	49.61	2.85	1.05
2004	51.72	2.44	0.65	52.98	5.21		49.33	3.32	0.25	50.69	2.27	1.05	5.08	2.72	1.10
2006	51.91	2.14	0.63	51.42	4.22		48.34	3.64	0.46	50.07	2.71	1.00	49.91	2.29	1.33
2008	51.97	2.08	0.41	50.43	5.82		47.90	3.32	0.29	49.90	2.86	1.04	50.06	2.86	1.07

注：三个指标都是根据劳动力市场指标和总人口计算得出的。根据表4的定义，表中的三列包含每个国家的下列指标：经济不活跃人口比例， $b = B/Q$ ；失业率， $u = U/Q$ ；空缺岗位率， $v = V/Q$ 。这三个劳动力市场指标都根据总人口数进行了调整，以进行更好的比较。表中只显示了整个时间序列数据的部分观测值。

资料来源：经济不活跃人口数、失业人口数和总人口数来自 ILO (2010)；空缺岗位数来自 OECD (2010)；英国和美国关于空缺岗位的时间序列数据来自 the Office for National Statistics (2010) 和 the Bureau of Labor Statistics (2010)。

可以看到，表6和图9都支持上文的论点：

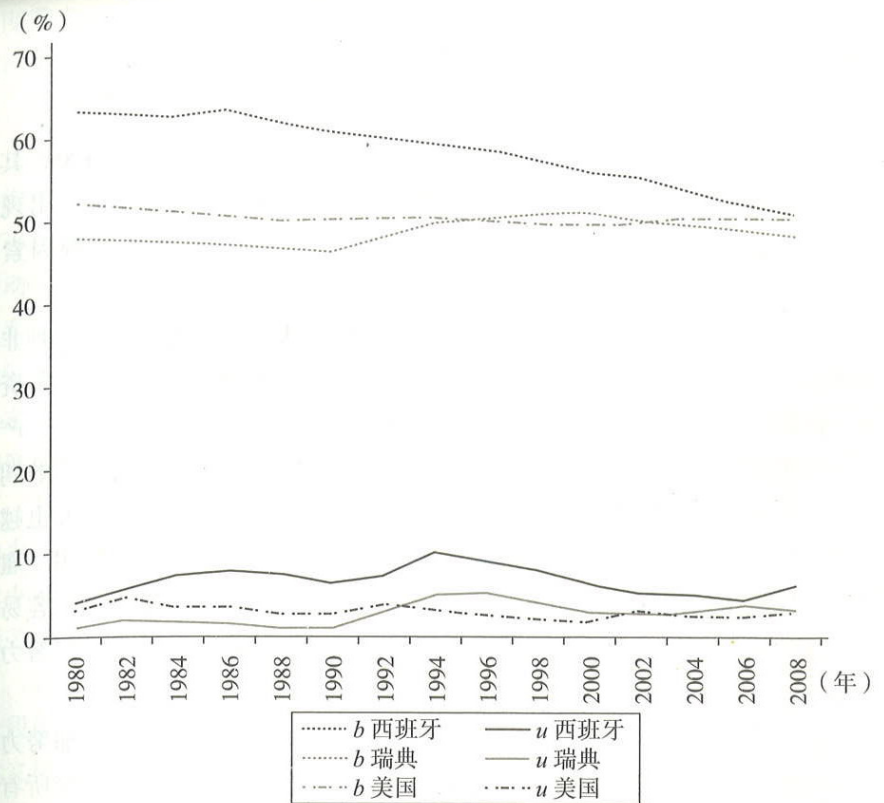


图9 三个老牌资本主义国家的经济不活跃人口比例和失业率 (1989~2009年)

注：根据表4的定义，西班牙、瑞典和美国的指标  $b$  和  $u$  的计算办法如下：经济不活跃人口比例， $b = B/Q$ ；失业率， $u = U/Q$ 。

资料来源：同表6。

- (1) 短缺和过剩到处并存。
- (2) 在传统的资本主义国家，过剩现象远比短缺现象普遍。<sup>①</sup> 瑞典等率先建立福利制度、试图降低失业率和增加经济活跃人群的国家也同样如此。

<sup>①</sup> 图6清楚地显示了影响经济参与率的两个重要因素：经济发展水平和制度的影响。表5的时间序列数据生动地反映了经济增长的影响。除劳动力市场制度的改变外，还有其他同时发生的因素在影响若干国家的参与率（经济不活跃人群的比例下降）。因此，如果不做补充分析，不能简单地看图6和表5的数据进行对比。



## 效率工资

另外一组导致持续产生失业的现象，与效率工资这一关键概念有关，其含义是雇主愿意向雇员支付高于“市场出清”水平的工资。效率工资的出现不是雇主的利他主义的表现，而是出于自身的利益考虑，它受到几种因素的影响。

首先是对怠工的预防。<sup>①</sup> 如果你要给每位雇员都配备一名监工，那是非常高的成本。因此，有必要把进行监督的成本与提供高于市场出清水平的工资成本做个比较，后者在很多情况下更为合算，尤其是在在工作业绩难以量化、产品质量十分重要的情况下。某种工作的智力投入越多，对员工的智力表现的期望越高，就越不适合采用计件工资或计时工资来激励，对怠工现象的监督也越困难。相反，必须设计内在的激励机制，这就是效率工资所能发挥的作用。雇主提供的工资承诺如果高于市场出清的工资水平，他在劳动力市场就能更容易地找到合适的职员。同时，雇员一旦获得了报酬较高的职位，往往也希望努力保留下去。

如果员工不仅担心失去高工资，还担心彻底丢掉工作，他们会更加努力工作，使效率工资的作用更为显著。确保劳动力处于长期过剩状态，符合所有雇主的共同利益，因为这将成为所有职场都能使用的武器。所以我们看到，即使那些最开明和最有社会责任感的雇主对于大力减少失业也不是全心全意地支持。他们当然不希望劳动力市场的形势逆转，出现短缺。于是，对于丢掉饭碗的担忧会时刻萦绕所有的雇员。这种情况甚至还反映在夏皮罗与斯蒂格利茨的一篇著名论文的标题上：《以均衡失业作为约束员工的工具》。<sup>②</sup> 上面的论述听起来非常像“马克思主义”，但实际情况的确如此。

在宏观经济学中，也存在著名的反对过于激进地消灭失业，尤其是反对“过度就业”的论点。其逻辑是，如果增加工资的动力过强，那么对劳动力的

<sup>①</sup> 效率工资理论的开创论文参见 Shapiro 和 Stiglitz (1984)，还有很多描述和因果分析方面的研究成果，包括 Milgrom 和 Roberts (1992，第 165~195 页)。

<sup>②</sup> 该机制发挥作用的时候，并不需要雇主们都意识到维持失业符合他们的共同利益。如果劳动力市场很紧张，就业率接近 100%，雇主们需要支付的效率工资将上升到无法承受的水平。当个别雇员在自己的工资谈判中提高要求时，他们也会在无意中把就业率压低到 100% 以下。

过度需求将推高成本和价格，引发螺旋式通货膨胀。这是一个合理的说法。<sup>①</sup> 然而它并不能否定其他机制的作用：维持失业现象，以此作为对劳动力的约束，符合雇主们的共同利益。

在讨论效率工资理论的时候，那些曾读过马克思的“产业工人后备军”理论和《资本论》第一部的其他内容的人应该感觉很熟悉 (*Capital*, Vol. I, 1978, 第 25 章)。后备军或者说过剩劳动力的存在的确有助于压低工资，加强劳动纪律，在雇主们希望扩大生产的时候提供可用的人力资源。我也不否认，效率工资理论与现代的社会主义和无政府主义运动的思想有类似之处，他们认为资本家通过支付更高的薪水，收买了一批“工人贵族”的效忠。<sup>②</sup> 学术道德要求我们，即便那些极不情愿甚至感觉耻辱的经济学家，也要承认这种思想上的亲缘关系（实际上还是首创权）。

劳动力过剩还有一个有利的影响：它的存在使调整变得更有弹性。如果劳动力的动员更加容易，生产的新建或改组也会更加方便。当然，这只是一种技术专家式的视角。闲置的产能和存货毕竟是死的，存放起来不会有多大损失，但劳动力的储备是人，让他们长期等待是一种煎熬。

我们需要暂停一下，回顾第 4 章的第 1 节所阐述的内容。在达到中等发展阶段的社会主义经济体中，出现了长期劳动力短缺，包括前苏联和东欧的社会主义国家。人们普遍认识到这种外部环境对劳动纪律的侵蚀。员工们有充分理由认为，即使不努力工作甚至完全消极怠工，他们也不会被解雇。就算被解雇，他们也能轻易地找到另外的工作。所有的经理人和监工都在不断地抱怨这种现象。许多员工出于基本的诚实品质和对工作的认同仍在忠实履行自己的义务，在有些地方，如果能有效完成任务，还可能获得物质奖励。然而他们从来不会被失业的幽灵困扰。我们也必须承认，这样的环境对雇员也有一定好处。参与就业给人以安全感，有保障的就业就更是如此。从劳动力短缺跳跃到劳动力过剩对社会整体而言是利有弊，而且这些利弊在各个社会成员中的分配是不均等的。

现在做个小结。持续地产生劳动力过剩、长期失业和就业不足是资本主义制度令人痛苦的固有特征，它给失业者带来了物质上的损失和精神上的羞辱，

<sup>①</sup> 这一关系是“自然失业率”理论的核心，其开创性研究参见 Phelps (1968) 和 Friedman (1968)，在任何一本宏观经济学教科书中都有介绍。

<sup>②</sup> “工人贵族”这一说法由巴枯宁首创，考茨基和列宁的著作中都有借用。

也带给读者带来了失业的担忧。雇用哪些人和解雇哪些人的决定，经常隐秘地掺杂着种族主义和其他形式的歧视。

一方面要接受和支持资本主义，另一方面又要求或承诺实现“充分就业”，那是相当不切实际的。在政坛和研究领域，许多人都想玩弄“充分”这个词，其实毫无意义。对我而言，只有100%才是真正的充分，而不是97%或93%。不仅是严谨的研究人员应该这样理解，许多利害相关的人同样如此。一个外科手术只有在从无一失败的时候，其成功率才可能是100%。医生或研究人员可以根据统计数据的结果宣布，97%的治愈率就能被认为是“绝对有效”，对药品而言已经是很令人满意的结果。但事实上，剩下的那3%的人却没有获得成功的治疗，将长期患病、残疾甚至走向死亡。

对于失业率来说，下降到3%可以被认为是宏观经济政策的巨大成就，但身陷3%之中的人们却不会因为“结构性失业”导致他们没有工作的说法而感到安慰。我们有可能也应该制定出更好的宏观经济政策和制度体系，来降低失业率，使其更能令人接受。这是一个现实的任务，值得为之奋斗。但我认为，宣扬任何一种政治制度能取得“充分”就业的成绩，都是在误导。

## 5. 对实证描述和因果关系分析的总结

### “均衡”概念的实用性

阅读本文的经济学家们现在可能已意识到，我一直在尽量避免使用“均衡”这个术语。提到它的时候主要是因为我在引用主流经济学的观点。

很少有其他术语像“均衡”那样引起极大的混乱。不同的学派对这个术语给出了不同的解释和判断。当一个经济体或其中的某个部分实现均衡时，有些学派感到欢欣鼓舞，有些无动于衷，还有些甚至深感惋惜。这样的讨论可以残酷地称为“聋子间的对话”，因为大家的阐述都不理会其他人的观点，或者会对“均衡”、“非均衡”、“反均衡”和“不均衡”（equilibrium, disequilibrium, antiequilibrium and non-equilibrium）等词汇发生严重的误解。

我对此不抱任何幻想。我不相信本文及其对均衡问题的有限思考能够使上

述的概念混沌恢复秩序。如果读者们能清楚理解我本人在写作时对这个概念的解释，以及我如何看待其实用性，那我就很满足了。<sup>①</sup>

“均衡”一词的拉丁起源来自“天平”：如果两边的质量相同，天平虽然可能发生摇摆，但最终会静止下来。这就产生了均衡一词的通俗含义：在一个系统中，作用力方向相反而大小相同，使得整个系统保持静止。这样一个广泛的含义适合用来描述各种系统的状态，没有什么疑义。

在数学模型中，均衡的概念可以用严格的条件来描述。我们不需要去寻找一种普遍的定义。在许多模型中，均衡的概念可以被解释和定义，还可以发现均衡的某些特征。可是，不同的模型所采用的是不同的定义。在分析动态系统的数学模型中，有可能发现“稳定状态”或者“固定点”，它们是与均衡类似的概念。但是也存在一些数学模型，不可能在其中给出均衡的定义。

让我们回到对现实问题的研究案例。均衡的概念被物理学家、化学家和生物学家们广泛采用，但我们还是主要看经济学领域。在现实社会的经济系统中，某些现象会在一定程度上表现出均衡的特征（完美或接近完美的均衡）。例如，一个国家或企业的账本会显示，在给定时期内，收入等于支出；一家仓库的发货可能正好等于进货等。

我们把问题聚焦：生产和消费的协调、买方和卖方的协调以及资源的配置，是属于什么情况？执行协调和配置的特定机能的市場表现为什么状态？那里是否存在某种均衡？我的回答是响亮的：不存在。在现实的市場中，通过现实的协调和配置机制的运动，没有也不可能有某种稳定的状态。竞争和冲突的各种力量都在持续变化之中。事实上，变化才是本质。如果偶然出现静止，那也是无聊和无趣的事情。

我既不是第一个，也不是唯一一个做这种阐述的经济学家。这是奥地利学派的核心论点之一，从米赛斯和哈耶克到柯兹纳，再到今天的继承者们。<sup>②</sup> 尽管在其他一些重要问题上存在分歧，但在这个极其关键的论点上，我毫无保留地和他们站在一起。

<sup>①</sup> 我强调说“本人在写作时”，也是一种自我调侃。因为我得承认，我自己对这个重要概念的认识也多次发生改变，如果这会让读者感到困惑的话，那只能表示遗憾。本文不会讲述本人的思想发展过程，也许以后会有机会。这里只是介绍我目前的认识。

<sup>②</sup> 奥地利学派的多篇开创性论文都非常重视这方面的思想，如：Hayek（1948），Kirzner（1973），Lachmann（1976），Lavoie（1985），Cowan和Rizzo（1996）。

利用家庭和市场上使用的老式天平，把食品放在一端的秤盘上，把砝码放在另一端，天平会在一段时间内来回摇摆。假如放置的砝码质量合适，摩擦力会很快让摆动缓慢下来，使天平恢复静止。此时就实现了严格意义上的均衡，因为摆动的停止表明两端的质量相同。然而，经济生活的现实与天平有着本质的不同。上文的第2章曾试图阐述这个问题，技术进步在持续发生，今年的供给与今天的需求不但与去年或者10年以前的情况在数量上不同，更存在质量上的差异。如果必须找出哪位学者的观点来补充奥地利学派有关市场的永恒变化的论述，那应该是熊彼特。持续的创新过程还意味着“市场均衡”这个词很难得到解释。如果采用实证研究而非规范研究的办法，那问题就不在于“市场均衡”是否有用。不管我们是否愿意，在现实的市场环境中都不存在均衡，也永远不可能出现。<sup>①</sup>

如果这个论点成立，那我们就没有必要区分下面这些不同说法：“完全必要”或均衡的产能储备，和“过剩”的产能储备；完全必要或均衡的存货，和过剩的存货；完全必要或均衡的劳动力储备，和过剩的劳动力储备。这样做不是无知，我不采用如此的区分是因为它们在现实世界中并不存在。

但这并不意味着均衡的概念在对现实市场的实证分析中毫无作用。这个概念可以用于假想实验，作为参照点。请看下面两个例子。在微观经济学理论中，我们可以借鉴瓦尔拉斯、阿罗和德布鲁以及他们的学术继承者们有关瓦尔拉斯均衡的决定条件的理论。瓦尔拉斯均衡的实现条件包括：在一个买方和卖方都掌握完全信息的市场，未来可以预测，任何调整都在瞬间完成，没有任何摩擦等。然后，我们可以分析现实的市场环境与这样的理想状态有哪些差距。

另一个例子是纽曼设计的增长模型，其中不存在技术进步，每个部分都以同样的速度在均衡路径上增长。我们可以观察，这种假设与每个经济体的实际增长表现有哪些不同，为什么有的产业会消失，有的会出现，为什么各个部分的比例在不停地发生变化。不过，一个抽象的数学模型只是针对虚拟世界的分

<sup>①</sup> 很明显，卡尔多正是因此认为，均衡经济学理论“与讨论的问题无关”（Kaldor, 1972）。我对此是认同的，但我还必须指出，从今天的立场来看，卡尔多当时被很多理论上的偏离激怒，以致丧失了客观判断，他拒绝承认有关均衡的数学模型可以在理论研究中发挥作用。卡尔多愤怒地撰写文章时所犯的错误，我自己也遇到过。我在自传里描述过当时的观点并进行了自我批评（Kornai, 2006, Chapter 10）。

析工具，纽曼均衡路径当然不能被视为对现实经济增长的描述。

我感觉有疑问的地方是，均衡这个概念被描述为系统成员处于“静止”状态。按照这一思路，如果成员们出于自身的利益，不再愿意离开均衡点，这个系统就达到了均衡。然而，该逻辑很容易导致概念上的同义反复。<sup>①</sup>

例如，不幸处于短缺经济中的买家没有办法在自己需要的时候、需要的地点买到称心如意的某种商品，最后，作为不得已的调整，他们购买了能够找到的某种替代品。他们对此只能自我安慰，因为大家都知道继续寻找的结果可能更糟。这种状态也可以被命名为一种静止状态、均衡状态，但这个术语却容易造成误导，而且是彻底的颠倒。

按照同样的思路，我对“失业均衡”这个概念也深感不安。我承认，在模型框架内，劳动力市场上会出现所有参与者都无可奈何的失业均衡状态，但如果那些本来愿意从事工作，只是由于某种就业障碍（上一章有过详细讨论）而未能如愿的人，又怎能高兴起来？人们担心更坏的遭遇，从而接受某种结果，把这当成静止状态或均衡的定义，其心理和社会心理基础是相当脆弱的。

从马歇尔和瓦尔拉斯到今天的主流学派，在所有经济学家眼里，高度重视均衡概念的定义和地位的榜样就是牛顿的经典物理学。但在我看来，如果需要从自然科学那里寻找灵感，那我愿与其他某些经济学家一样选择达尔文的生物进化学。自然选择和生物进化与市场经济中的增长和技术进步有着发人深思的相似性。

每个新企业和新发明都可以视为随机产生的变异。新产品、新技术、新的组织方式和新的企业机构都在为生存而战。有些能获得成功，不但存活下来，还能繁衍壮大，发明被广泛采用，企业扩充规模。有些变异则时运不济，发明没有被实际应用，企业倒闭关门。这也是一种自然选择。生与死的决定不是来自一个具有理性、远见和无上智慧的中央意志，而是取决于关键的生存力指标。运气也起着很大作用。总体而言，这个机制运转良好，但也并非一贯如此：有些失败的企业本来很有价值，而有些存活时间很长的企业却本可以被更好的企业取代。

<sup>①</sup> 从预期角度对均衡概念进行定义的理论对此产生了重要影响。均衡状态被描述为理性预期的实现。本文没有讨论这方面的问题。