



■ 张维迎 北京大学国家发展研究院教授

1、纳什和博弈论如何变革了经济学？

日前一位重要的博弈论大师约翰·纳什去世了，由此引起了很多人对博弈论的更大兴趣。之所以用《博弈与社会》这个题目是因为没想到更合适的，我今天要讲的主要内容是简单给大家概括一下我认为有意思的东西。当然，一个多小时的时间也不可能讲太多，而且我希望能够给大家留些时间去提问题。

首先我们来看经济学由于纳什或纳什均衡发生的革命性变化。在古希腊时代,经济学是道德哲学的一部分。我们知道,亚当·斯密创立了现代经济学,但他的第一本书是《道德情操论》,斯密本人是格拉斯哥大学的伦理学教授。在斯密之后,纳什之前,经济学被简单地理解为有关物质财富的生产和配置的理论。

在纳什之后, 经济学研究扩展到了社会制度中的激励理论, 或者简单说经济学研究的是制度和激励问题。经济学关注资源配置理论的时候, 主要的研究对象就是财富如何生产, 如何分配。当经济学关心一般的社会制度和激励机制时, 关注的核心就是人与人之间的相互合作。在这样一个转变中, 纳什均衡的提出发挥了革命性作用。

我们再往前看以资源配置理论为核心的传统经济学, 这个经济学的核心就是价格理论。我们过去说, 懂不懂经济学就看你懂不懂价格理论。比亚当·斯密更早的人, 如威廉·配第(William Petty)这些人研究经济学, 他们更喜欢用数学方式, 用线性代数的方式分析问题。用这种方式分析问题最容易出现的就是物质财富, 因为物质财富在物理上有一个量, 从货币的角度看也有一个量。

物质财富生产特别适用于定量分析法, 所以新古典经济学也越来越转向研究物质财富的生产和分配。在这个理论中, 每个人决策时面临的是非人格化参数。我在决定买什么不买什么, 或者找这个工作还是那个工作时, 我有一个偏好, 剩下就是各种参数, 比如价格参数, 苹果或桃的价格啦, 工资的多少啦, 我不需要考虑其他人的想法。

但是我们知道, 社会关系是人与人之间的关系。在新古典经济学中, 我们假定市场中, 人与人之间的关系完全通过价格来体现, 或者说价格是一个参数, 对所有人都一样, 每个消费者都有自己的最优选择, 然后就形成需求函数; 每个生产者有自己利润最大化的选择, 于是就形成供给函数。在市场当中, 似乎总有一只无形之手来让需求和供给相等, 于是达到了所谓的均衡, 这就是传统经济学的基本理论。

所谓均衡, 在传统经济学里是指价格处于这样的状态: 需求等于供给, 每个消费者达到效用最大, 每个生产者达到利润最大。在均衡状态下是没有经济利润的, 所有的收益都是成本。如果大家记得经济学思维中的成本是指机会成本的话, 均衡情况下就没有利润。这是我们传统的经济学思维。这种思维进一步简化就变成了成本-收益分析, 或者叫供求分析。萨缪尔森的书里有一句话: “只要会会需求和供给, 鹦鹉都能变成经济学家。”这是传统的经济学。

经济学家在发展出了这套非常成熟的价格理论之后, 做出的数学模型确实非常完美, 非常漂亮。将这些理论应用于分析其他社会问题, 我们一般叫做理性选择理论。但当我们这样去分析社会问题时, 就会面临很多困难。其中一个困难就是大量的经济行为其实是没价格的。最简单的例子, 政治谈判和竞选就没有价格, 这里的没有价格是指没有用货币表示的价格。

另外, 人在实际行为中关心的不仅仅是物质利益, 比如我找工作并不是只关心工资, 我还关心其他的东西, 如工作环境、对我未来职业选择的影响, 还有这个职业的社会声誉, 等等。有些职业的工资可能不高, 但是它使我们很有成就感, 比如媒体。媒体的工资不高, 但是为什么好多人都愿意做呢? 它有着很高的成就感, 就是我们通过愿意做每天有着很的事情, 媒体人往往有另外一种追求, 而不只是为了赚钱。经济学家其实也是这样。这个是传统经济学难以分析的。

还有, 人在决策当中, 面临的不一定只是非人格化的参数, 还面临其他人的决策。有时候, 我做还是不做某件事, 依赖于你做还是不做; 或者我愿意选择什么, 依赖于你会选择什么, 而不是简单地只面临一个价格或收入等非人格化参数。

举一个例子, 两个人讨价还价, 在这个过程中, 我出什么样的策略, 我的价码是多少, 依赖于我判断你的价码是多少, 而不是说我通过一个确定的价格参数来机械性地选择。即

好的纳什均衡和坏的纳什均衡

便应用于最简单的例子——去地摊上买东西, 传统价格理论的局限还是很大。

传统经济学假定, 一个生产者在决策时, 首先要要素价格给定, 产品价格给定, 然后最小化成本, 选择最优生产组合, 再选择最优产量。假定中产品价格是给定的, 但竞争中并非如此。

比如寡头企业决策时, 价格不是给定的, 价格多少不仅取决于他, 还取决于于对方。这时传统的价格理论就没有办法分析。这点也说明, 为什么我们现在要把寡头竞争的博弈论模型追溯到年法国数学家古诺 1838 年做出的分析, 他是第一个真正用数学模型分析寡头竞争的人, 而局限于一个我们所谓的完全竞争市场。

当然了, 这些局限性并没有阻碍经济学家试图把传统价格理论应用于非经济问题的研究, 这方面的典型是诺贝尔经济学奖得主、芝加哥大学教授加里·贝克尔(Gary Becker), 他试图用经济学分析方法去分析所有的人类行为问题。另外我们知道, 法经济学也是试图用成本—收益方法来阐述法律制度。

现在来看, 这些研究总体上并不令人满意, 这时就迫切需要一种新的更为一般化的分析方法来克服刚才讲的价格理论的缺陷: 它可以分析非价格、非物质的东西, 可以分析人与人之间相互关系或互动情况, 这个理论就是博弈论。

在纳什之前, 也有其他的经济学家提出过类似理论, 比较有名的像刚才讲的古诺, 像计算机的发明人冯·诺依曼, 还有他的合作者摩根斯特恩(Oskar Morgenstern), 但博弈论真正成型还要归功于纳什。所以一般我们认为, 非合作博弈的创始人就是纳什, 他把之前已经发展出来的理论组装起来。纳什对博弈论的贡献, 某种意义上就像亚当·斯密对经济学的贡献, 在斯密之前也有好多经济学理论, 亚当·斯密第一个把它们组成成系统理论。

我们知道, 纳什发展出的最重要的概念就是纳什均衡。我在这稍微给大家介绍一下。首先我要纠正一点, 一般我们讲的博弈论就是非合作博弈论, 但在中文里特别容易误解, 让人觉得非合作博弈是研究人怎么不合作的。事实不是这样, 非合作博弈研究的是每个人独立决策的结果会是什么样。我们恰恰是希望用非合作博弈理论来解释人们为什么不合作, 只有搞明白了为什么不合作, 才能更好促进合作。

先来给大家做一个简单概述, 什么叫博弈论。博弈论指的是研究人与人之间行为互动的一般理论。所谓社会, 就是互动。很多经济学家对博弈论的应用范围评价极高, 比如诺贝尔奖得主奥曼(Robert Aumann)。另外一个经济学家哈特(Sergiu Hart)在文章里说, 博弈论可以视为整个社会科学理性一脉的总括。我们研究的人的行为有理性的有非理性的, 博弈论为理性行为分析提供了一个统一场理论。

这里我们讲的社会, 可以广泛地理解, 既包括人类个体组成的社会, 也包括其他各种个体组成的群体, 比如公司、国家、动物、植物、计算机这些都可以应用博弈论分析。我们研究国家之间的关系, 研究中央政府和地方政府的关系, 都符合我们讲的社会范畴。

博弈论不像经济学或政治科学等学科其他分析工具那样, 采取不同的就事论事的框架, 来对各种具体问题进行分析, 如完全竞争、寡头、垄断、国际贸易、税收、选举、遏制, 还有动物行为等等; 相反, 博弈论是先提出一些一般的原理, 这些原理适用于所有互动的情况, 然后考察这些方法具体应用会产生何种结果。

博弈论为我们分析整个社会变化和社会结构提供了一个微观基础。曾获得诺贝尔经济学奖的迈尔森(Roger Myerson)教授曾说, 发现纳什均衡的意义可以和生命科学中发现DNA 双螺旋结构的意义相媲美, 纳什构造了经济分析新语言的基本词汇。另一位经济学家这样认为, 越来越多地, 博弈论语言变成了经济学语言, 纯理论指的就是博弈论, 我们现在讲的社会科学的纯理论就是指博弈论。这种评价是相当高的, 但并没有说过其实。

有了这样的理论, 所以在纳什之后, 我们的经济分析不再是简单的资源配置理论, 不再是简单的价格理论, 而是可以分析各种各样的制度的理论, 包括市场制度。传统经济学只分析市场制度, 我们现在要分析大量的非市场制度; 传统的经济学只分析物质生产和分配, 我们现在不仅分析物质的, 也分析非物质的, 不仅分析经济问题, 也分析社会、政治、文化问题以及它们之间的相互关系。我们还可以分析制度是怎么演化的, 这个传统经济学没有办法分析, 有了博弈论之后, 制度演进分析变得容易。

所以博弈论使得经济学发生了根本性转型, 也正在使得其他社会科学发生这种转型, 包括政治科学, 包括法律, 甚至包括最基础的心理学, 社会学等等, 包括对动物的研究, 动物学, 都在发生一些重要的变化。

我讲的一个基本问题是, 从博弈论或者以纳什划分为界, 纳什之前和纳什之后, 经济学是很不一样的。现在大学教科书里边, 基本的经济学原理主要是资源配置理论或价格理论, 这个过去叫微观经济学。博弈论作为单拎

出的一部分, 有些教科书会把它放进去, 但是并不是所有都放进去, 放进去的分量非常有限, 比如曼昆(Gregory Mankiw)的教科书《经济学原理(微观部分)》》有一点博弈论的内容, 但是大部分内容仍然是以传统价格理论为主。

2、纳什均衡的制度含义

接下来稍微给大家介绍一下纳什均衡, 为什么这个均衡如此重要, 如此有意义。简单地说, 纳什均衡是一种分析和预测互动情况下理性人行为的概念工具, 我们假定人们是理性的, 然后在相互作用的情况下, 研究他们会有什么样的行为和后果。

我们先从预测人的行为说起。在传统经济学的价格理论里, 一个人的效用函数依赖于自己的选择, 比如你买了几个苹果, 买了几个梨, 对你的效用产生了怎样的影响。在博弈论里, 一开始就改变了这个效用函数, 一般讲我们消费者的偏好, 不仅依赖于你的行为, 而且依赖于别人的行为, 所以在这里我们看到的跟传统经济学不一样, 第一要看到自己的行为, 第二还要看别人的行为, 别人可能是一个人, 也可能是好多人, 可能是 N 个人, 这是非常重要的一点。

我们假定每个人都是独立做出决策, 然后来研究什么样的结果最可能出现。纳什均衡就是这样最可能出现的结果, 或者说最可能出现的一种策略或者策略组合, 每个人都根据自己情况选做出最优选择, 如果一个策略或者一个行动选择变成纳什均衡, 意味着什么? 意味着给定别人的选择, 没有任何人有积极性改变自己的选择。举例来说, 今天晚上在这里, 每个人来的时候都要选择自己的座位, 我们现在都已经各自坐好, 你可以说我们现在就是纳什均衡, 因为没有人有积极性跑到另外一个地方去。

现在我们把纳什均衡理解为一种制度安排来思考一下。如果一个制度是纳什均衡意味着什么? 意味着所有人都要遵守它; 反过来说, 如果一个制度不是纳什均衡, 至少有一部分人不会遵守它, 也可能所有人都不遵守它。这就是为什么纳什均衡对我们分析社会制度和制度演化如此重要的一个原因。我这样讲比较抽象, 下面我们会用简单的例子说明这点。

首先, 纳什均衡是没有价值判断的, 它不告诉你这个制度安排是好还是不好, 它只告诉你如果每个人都选择自己认为的最优行为, 会造成什么样的结果。当然了, 如果我们主观上给纳什均衡本身加一个价值判断的话, 可能有好的纳什均衡, 也有不好的纳什均衡。什么叫好的纳什均衡什么叫不好的纳什均衡, 依赖于我们用什么标准衡量。传统的经济学里我们会用社会福利函数, 或者帕累托效率标准来衡量, 如果这个制度达不到帕累托均衡, 我们说这是不好的, 达到了就是好的。

我想, 衡量好还是不好, 更一般的标准就是合作。为什么这样讲呢? 因为人类的所有进步, 每个人利益的改善都来自相互的合作。所以什么叫好的纳什均衡? 如果这个纳什均衡达到了一个合作的目的, 合作是纳什均衡, 我们说这就是好的纳什均衡, 否则就是坏的纳什均衡。我们先从坏的纳什均衡讲起, 如果学了一点博弈论就知道, 这是教科书上常提的囚徒困境。为了下面要讲的内容, 我稍微为大家复习一下。

假如两个人组成的社会, 每个人都有两个选择, 可以选择合作, 也可以选择非合作。当然了, 这个概念可以作为一般应用, 也可以理解为在团队中的努力还是不努力, 合作是努力, 非合作是不努力。你也可以讲合作就是诚实, 非合作就是欺骗。为了方便简洁我们简单地用合作和非合作。这样的话, 每个人最终会得到什么, 不仅依赖于你选择合作或不合作, 还依赖于别人的选择。

在两个人情况下, 有四种可能性: 第一种两人都合作; 第二种两人都非合作; 第三种是第一个人合作, 第二个人不合作; 第四种第一个不合作, 第二个人合作。对应这四种可能性, 每个人的报酬, 或者说叫效用, 或者博弈论中叫支付, 是不一样的。我们先从个人讲起, 因为博弈论是从个人的决策开始的, 先要研究对单个人来讲, 什么是最好的。

在这个案例里面, 对单个人来说最好就是别人合作, 他不合作, 他就赚便宜了。比如别人好好干活, 你偷懒, 你分享别人的成果, 你赚了; 其次是两人都合作; 再其次是两个人都不合作; 最糟糕的是你合作, 别人不合作, 你被别人骗了。凡是满足这样一个条件的, 都叫囚徒困境。

图一：囚徒困境

		B	
		合作	不合作
A	合作	3, 3	-1, 4
	不合作	4, -1	0, 0

图一用具体的数值描述了这个博弈。在这里, 从 A 的角度来看, 如果 B 合作, 他不合作, 他得到 4; 其次是两人都合作, 他得到 3; 再其次是两人都非合作, 他得到 0; 最糟糕是什么? 他合作, B 不合作, 结果他得到-

1。图一里边的第一个数字代表第一个人的报酬, 第二个代表第二个人的报酬, 这是表述的惯例。

我们看到, 如果是这样情况, 每个人出于自己的利益选择, 他会选择什么呢? 从 A 的角度出发, 如果 B 合作, 他最好是合作, 他得到 4; 如果 B 不合作, 他合作的话就亏了, 不合作是 0, 所以从他的利益出发, 最好选择不合作。从 B 的角度讲也是一样, 所以每个人的最优选择都是不合作, 结果是每个人都得到 0。而我们作为局外人, 明明看到这样更好: 如果两人都合作得到 3。这就是我们过去讲的两个人理性与集体理性的矛盾, 这就是囚徒困境的意思。

两人都非合作的纳什均衡显然是坏的均衡, 从这里我们就可以体会到制度的意义。一个制度其实是某种游戏规则, 实际上就是博弈规则, 这就是制度的含义。它不仅在影响人的行为, 还决定着什么样的纳什均衡会出现。什么叫好的制度? 好的制度就是能使得合作变成纳什均衡的制度; 反过来, 坏的制度就是使得不合作变成纳什均衡的制度。在刚才这个例子当中, 我们可以设想这也许就是人民公社制度, 大家吃大锅饭, 谁好好干活谁吃亏, 谁耍奸谁得到好处, 最后结果是大家没饭吃。

图二：包产到户下的纳什均衡

		B	
		合作	不合作
A	合作	3, 3	-1, 4
	不合作	4, -1	0, 0

我们现在看制度改变起什么意义? 假如人民公社变成包产到户, 在包产到户的情况下, 你赚不了别人的便宜, 你只能吃自己地上打的粮食。以前, 如果农民两人都合作, 每人可以得到 3, 俩人都非合作, 得到 0。现在就变化了, 如果第一个农民 A 努力, 第二个农民偷懒的话, 农民 A 自家能拿到 3, B 只能拿到 0; 反过来说, 如果 B 努力, A 偷懒的话, A 只能得到 0, B 得到 3。你看, 在这样的制度下, 纳什均衡是大家都会合作。

从 A 来讲, 不管 B 合不合作, 我自己一定会好好干活, B 最后选择也是好好干活。结果就从前面坏的纳什均衡变成了好的纳什均衡, 即合作的纳什均衡, 这就是我们讲的制度的意义: 制度影响纳什均衡, 一个好的制度鼓励人们合作, 一个不好的制度使大家不合作。

但我们马上遇到一个问题了, 农业是由于生产落后, 不需要实际上物理上的合作, 每个人自己种自己的地就好。当然, 互助组合作社之类的我们先不讲它。但是, 任何一个工业企业不可能分成个体生产, 如果你把一个大企业分成一个个个体户的话, 公司的价值就没有了, 我们需要这个公司本身就是因为它可以创造出单个人创造不出的东西, 也就是说 N 个人放在一块创造的价值大于每个人创造的价值乘以 N, 这才是我们需要团队的原因。所以我刚才这个例子, 并不能完全概括我们面临的困境。

图三：私有产权下的纳什均衡

老板	雇员		
	合作	不合作	
	合作	3, 3	2, 0
	不合作	-1, 3	0, 0

如果在集体当中, 囚徒困境一定会存在的话, 怎么解决它? 现在来看, 私有产权就能解决它。

这里的私有产权可以这样设想, 还是一个博弈, 如果现在企业有一个老板, 一个雇员, 老板合作, 雇员也合作的话, 每个人得到 3; 如果老板合作, 雇员不合作的话, 老板就可以处罚雇员, 雇员只能得到 0, 老板还能得到 2; 如果雇员合作, 老板不合作的话, 老板照样付雇员 3 的工资, 老板就亏了。从这里可以发现, 一个人干活只能创造 2 的价值, 两个人干活可以创造 6 的价值, 于是企业的价值就出现了。这时候, 雇员和老板都会选择合作, 如图三所示。

所以, 在有了私有产权的博弈当中, 我们看到合作就变成一个纳什均衡了。为什么现代企业制度能够出现并且持续存在, 即便它有好多种如偷懒这些问题, 就是因为有私有产权制度, 简单说有老板。如果一个企业没有老板就会变成囚徒困境博弈, 纳什均衡就是没人合作, 这就是国有企业的问题。如果一个企业有老板了, 那就改变了制度, 改变了游戏规则, 合作就变成了纳什均衡。这就是前面讲的制度的意义。

再进一步, 一方面人们是在给定博弈规则的情况下做出选择, 我们可以预测纳什均衡是什么; 另一方面, 纳什均衡本身可以反过来分析游戏规则本身, 因为游戏规则也是人类博弈的结果, 它不是由一个外在力量给我们强加的, 哪怕计划经济也是人类博弈的结果。当然在这个结果当中, 可能我们发现我们犯了错误, 我们可能会改正它。我现在要跟大

家介绍一下, 好多情况下, 我们可以用纳什均衡理解制度本身, 甚至分析制度的演化。

17 世纪有本很重要的书——霍布斯(Thomas Hobbes)的《利维坦》, 《利维坦》描述了一个没有政府的自然世界里充斥着人与人之间的战争, 即所谓丛林规则。丛林规则就是我们刚才讲的囚徒困境。霍布斯认为, 为了走出丛林规则, 我们就需要一个政府, 有了这个政府以后, 我们有一个统一的意志, 强加统一的规则, 不能通过伤害别人得到好处, 结果对所有人都好。

霍布斯的这个解释从博弈论角度看, 就是通过设立政府使一个社会走出了囚徒困境, 变得可以合作, 这是他最早分析制度的含义。尽管他本身并没有用博弈论的概念, 但现在我们完全可以用纳什均衡的理论去理解它。

我们来把一般的制度框架概括一下。我们刚才讲的制度就是游戏规则, 我们分析制度要看制度是怎么产生的, 是怎么执行的, 这是两个最重要的维度。无论产生还是执行, 都有两种类型, 一个是分散的, 或者非专业化的; 另外一个是集中的、专业化的。分散的、非专业化的就是我们社会当中自发的从上到下, 集中的、专业化叫从下到上。

如果这样, 我们就看到有四种可能的制度, 第一种制度无论是产生还是执行, 都是分散出现的, 你不知道谁设计的它, 你也找不到有专门的人去执行它, 这个就叫做社会规范; 第二种, 无论它的产生还是执行, 都是有集中的、专业化的机构做的, 这个我们叫法律, 或者立法, 我们有专门的立法机关来造法, 在中国是全国人大, 也有专门的机构来执行它, 包括警察、法院。

但是一般的社会规范, 我们找不到谁制造了它, 我们不知道谁执行它, 在人类长期历史演化中每个人都可以执行, 每个人都可以不执行。你看到一些违反社会规范的行为, 你可以谴责, 你也可以睁一只眼闭一只眼。总之, 我们最重要的游戏规则, 就是社会规范和法律。

另外在介于这两者之间的还有其他类型, 一类的产生是自发的, 分散的, 非专业化的。但是它的执行有一定的专业化, 比如英国的习惯法。习惯法不是任何一个机构制定的, 但是它是法院执行的。

我们中国传统的“礼”, 可能介于社会规范和立法之间。“礼”不是政府制定的, 但另一方面, 大家觉得好像是儒家制定的“礼”, 儒家是一批专业化的人才。在礼的执行当中, 既有分散的, 如大量的“礼”是民间执行的, 也有一部分得到了法律认可, 比如县太爷在判案的时候, 传统的“礼”是判案的依据。所以它介于这法律和社会规范二者之间。

我们还可以更多地细分, 我们的社会制度或者游戏规则简单地说是这样一些东西, 都可以归结这几类, 包括我们的文化也可以说是社会规范。

3、社会规则的基本功能: 激励与协调

下面我简单给大家分析一下社会规则的基本功能。这里要讲的有些是在某种制度情况下人们如何行为, 有些是这些制度本身是如何形成的。我大体总结了人类所有的制度、游戏规则的四大作用: 第一是激励功能, 提供正确的激励, 当然也可能是错误的激励; 第二是协调功能; 第三是承诺功能; 第四是信号功能。

图四：作为激励机制的社会规范或法律制度

	合作		不合作	
	合作	不合作	合作	不合作
合作	3, 3	-1, 4-by	4-ax, -1	0, 0
	4-ax, -1	0, 0	0, 0	0, 0

先看一下激励功能。刚才已经谈到的私有财产制度就有激励功能, 包产到户就是激励机制。现在更一般的来讲, 在博弈当中, 如果没有这些游戏规则的话, 我们得到就是不好的纳什均衡。法律是来干什么的呢? 如图四, 假如说合作是一个合同, 我们大家都要合作, 如果你不合作的话, 你直接得到 4, 但这个时候你要受到一点惩罚, 就是 y(或 x), 从你的所得中减去, 因为你违反了合同, 要补偿人家。

我们先假定不去考虑图中的 a 和 b。只有这个惩罚 y 和 x 足够大, 人们才可能去合作。在这里, 只要 y 大于 1, x 大于 1, 当你预计你不合作之后, 带来的好处要扣除你要承担的惩罚, 你觉得不如合作的话, 你就会选择合作。这个时候这种法律就可能使得合作成为纳什均衡。

社会规范也一样, 社会规范不像法律那么有强制性, 但就像我刚才讲的, 我们人不仅重视物质的东西, 也重视非物质的东西, 当别人鄙视你的时候, 你会有不愉快的感觉。比如, 你看到别人的东西, 你抓过来自己吃, 人家不打你, 甚至骂都不骂你, 人家只瞪你一眼, 你会有什么感觉? 你心里会有点不舒服, 这就是一种成本。如果有了规范, 所有人都要遵守这个规范, 你违反这个规范的话, 就相当于你要付出一种成本。

(待续)