

# 讲义 1 和 2 — 介绍和初步应用：

## 最低工资的争论和原因的经济学推理

David Autor  
14.03 2004 秋季

### 1 课程简介

这是一门中级的微观经济学理论课程。我认为你们已经学习过了 14.01。而如果你们也学过一些计量或是统计学的话，会对学习这门课很有帮助。

这门课有三个主要内容：

1. 经济学理论 - 它讲的是什么？ 有什么用处？
2. 因果联系 - 我们怎么看待它？ 当我们研究的时候怎么发现它？
3. 经验的检验 - 经济学理论是一种掌握，预测和解释现实的方法，但这并不意味着不依靠数据。数据使我们能够检测理论的相关性和构建一个描述理论预测的准确程度的标尺（理论预测常常是方向性的(正的/负的)但很少定量地表示）。根据我们讨论的需要，我们会分析大量的实验和准实验。

定义： 准实验 — 一种实验方法，那些创造了与某个实验相类似的条件的事件可用来当作实验使用。例如：出生时就被分开的双胞胎：这个准实验可以被用于分析遗传—环境问题。而假设两亿人每人买一张彩票，其中会有 100 个人中奖。我们可以用这件事来探讨财富对于人们幸福，健康，离婚，肥胖等的影响。

为什么运用准实验的方法？因为许多关键的经济问题都围绕着人们生活中的重大选择及其结果：健康，财产，教育，风险等等。这些东西并不总是可以拿来做实验，所以我们在现实世界中寻找与我们所要求的实验条件相类似的偶然事件。如果在伦理上可行的话，我们就会进行实验。

#### 1.1 简要讨论：经济学方法论 - 或‘为什么是经济学理论’

##### 实证经济学

- 研究“是什么”。一种尽量不带价值判断的描述性行为。
- 建立模型以解释和抽象概括我们观察到的现象。

- 基于这些模型做出预测。

## 规范经济学

- 评估“应该做什么”。为经济政策提供建议。
- 有时实证经济学提供了足够的工具告诉我们一项政策优于另一项。（问：什么时候会发生这样的事？答：当一项政策帕累托优于另一项时。但是这样的情况并不多见。）
- 定义：帕累托改进 - 一种选择/政策/结果能至少让一个人的情况变好而同时不使其他人变坏。虽然对此并不是完全没有争议，但起码这是个谨慎的道德评判标准。
- 不论什么情况下，存在帕累托改进的政策选择是很罕见的（原因 - 理论表明人们已经做出了这样的改进）。大多数政策选择都有赖于价值判断，伦理偏好以及对相互抵触的目标的权衡等等（例如，就业和通货膨胀，平等和效率）。
- 尽管经济学理论不会告诉你应当选择什么样的政策，但它会有助于你进行取舍。

### 1.1.1 经济学方法相对于其他社会科学的长处

- 严格性：给定假设，运用科学的方法得出具有内在一致性的结论。
- 连贯性：建立在一些最重要的原则和理论基础之上。
- 思辨性：作出可验证的有力预测，其中许多显得很正确。
- 实用性：有助于你更好地理解世界怎样运行。

### 1.1.2 经济学方法的弱点

- “经济学一个显著的特点是它对个人行为 and 动机的分析有着令人吃惊的不成熟...” - Paul Krugman
- 例：简化的和过强的假设常常是不合适的，也不可能完全正确（比如，人们总是理性地追求自利 - 不同于自私 - 的目标）

### 1.1.3 然而缺陷中也有优点

- 我们有关于整个世界的模型。它就叫做世界 — 而运用它又过于复杂。
- 典型的经济学理论提供的是一个十分简化和高度程式化的世界，但是它对我们却相当有用。
- Friedman：“测试模型是否有效要看它预测真实经济现象的准确度，而不是它的假设的真实性”。
- Friedman：“言简意赅的假说才重要”。
- 我们的方法：简单的模型，深刻的洞察。

#### 1.1.4 三个对经济学方法的深刻的洞察

- “人们最大限度地发挥他们拥有的资源，实现效用最大化”。如果你长期地观察人类的行为，不管从实证上还是规范上你都会发现一件很令人惊奇的事：假定人们总是试图做出对自己最有利的选择，能在很大程度上解释人们交往中很多的现象。也要注意：其他的假设吸引力就显得小多了……
- 均衡 - 市场会整合所有个人的选择而产生集体性的结果，而这些结果往往与个人决策有极大的不同。
- 用效率作为标准来评价均衡的性质：
  - 面临抉择时，个人试图做出他们的最佳选择。
  - 对所有人来说，市场均衡是否是最好的结果（即，是否达到帕累托有效）？并没有明显的理由去做这样的假设 — 也就是假设中央统一计划并不比每个人出于自私独立的做出选择所得到的偶然的更好。
 然而，经济学中的一个非常好现象是：在一定条件下，市场能够产生所有可能的结果中最好的结果。  
而且这样的状况没有出现时，理论可以解释其中原因并指出如何得到更好的结果。

好的结果：福利经济学基本定理：自利导致有效率

坏的结果：外部性，市场失灵

## 2 课程要求和期望

1. 参考阅读 - 至少三分之一课时会用于讨论最近发表的经济学论文的部分细节。为了准备课堂讨论，你应该在课前读完这些文章。不要求理解文章的全部意思，但这些都是你应该弄懂的：问题，基本方法和关键结论。

2. 课堂参与 - 要求。（见教学大纲）
3. 六份习题集 - 除自动忽略的一份外，其他的五份在你的成绩中各占 5%。星期二下午五点之前交。迟交不算成绩 - 这就是制定自动忽略规则的原因。所以不要在学期伊始就浪费了你的这项权利。习题集中包含基于参考读物和讲座中的标准，正式的问题。
4. 三次考试，每次占 25%。两次随堂，一次在考试周。考试内容是前一次没考的东西。
5. 关于到课。很大一部分的资料不在教材上，特别是参考阅读材料以及我们所涉及到的一些理论资料。如果你不按时到课的话，你会感到难于完成习题和考试。而假若你打算只在考试的时候露面的话，我强烈建议你不要选这门课。
6. 不会严格的按照曲线来评分。每个人都可以做得很好（或差）。课堂参与加分在总成绩评定之后。如果你不多加努力的话，很有可能只能得到 C。而我若认为你只能得到 D 时，会劝告你最好放弃这门课程。但是在你放弃之后我也并不会给你帮助。
7. 课下支持：
  - 习题课在每个星期五，早上九点，十点各一节。课上你们的助教会解释课堂资料，帮助准备考试和评讲习题。头两节习题课 - 包括第一个星期五的 - 讲授 14.03 所需的数学工具。我在讲课时会用到它们但不会仔细的讲。因此请不要缺课。
  - 助教上班时间。每星期三下午两点到四点。星期三前交习题。
  - Autor 上班时间。星期二下午三点到四点我在办公室。
  - 对课堂主题和习题的疑问。使用课程网站，我会每天查看。不要发送实质性的与课程有关的 E-mail 给我们（个人问题可以通过 E-mail）。

### 3 第一个例子：最低工资和就业的争论

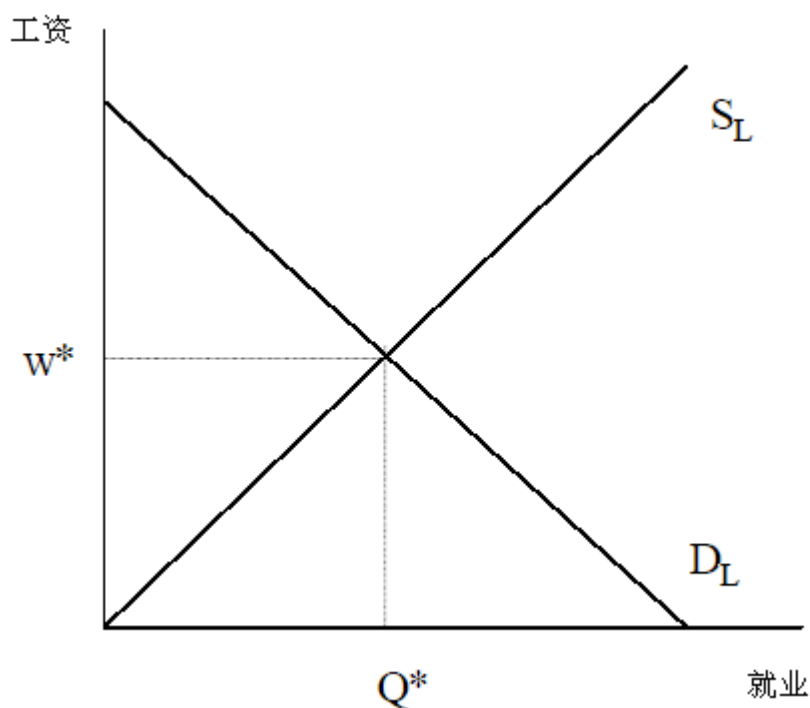
#### 教学日程

1. 教材中竞争劳动市场的模型
2. 教材模型中最低工资的影响
3. 模型的假设
4. 当我们去掉一项假设 - 如价格接受假设时，情况会发生什么变化？

5. 当雇主拥有市场势力时最低工资的影响
6. 检验教材的模型和选择
7. 经济学中的自然实验
8. Card 和 Krueger 的文章

最低工资：经济学中一个很老的话题，并且一直处于争议之中。

## 4 教材中工资和就业的模型



劳动供给曲线：劳动市场中所有潜在的工人，按照“保留工资”，即他们所愿意接受的最低工资（从低到高）就业排列。

劳动需求曲线：劳动市场中所有潜在的雇主，按照他们对工资的支付意愿来雇佣工人（从高到低）排列。

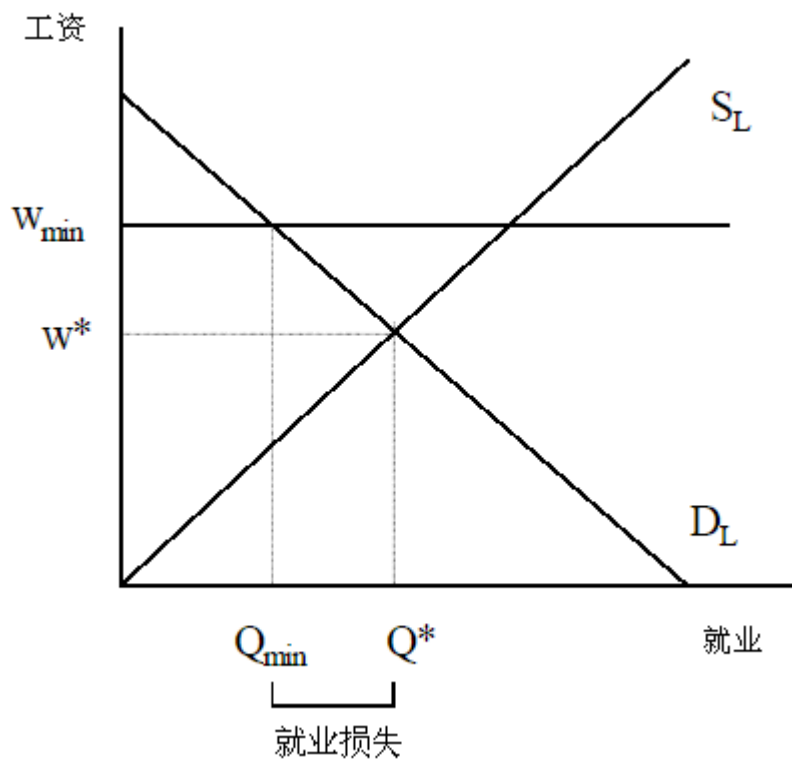
问：这个模型中的关键变量是什么？工资还是受雇的工人数？

答：都不是。它们都是被决定的内生变量。

**定义：**内生：内部决定的。同原因相对的结果。

**定义：**外生：外部给定的。作为原因的变量。

在这个劳动市场中，如果我们确定一个最低工资会有什么发生？



工资：

$$w_{\min} > w^*$$

就业：

$$Q_{\min} < Q^*$$

问：如果这个模型是正确的那我们为什么要设定最低工资标准？

一种答案：总收入

$$w_{\min} Q_{\min} \geq w^* Q^*$$

即使就业减少，所有工人的总收入也可能会增加。。

问：需要什么条件？

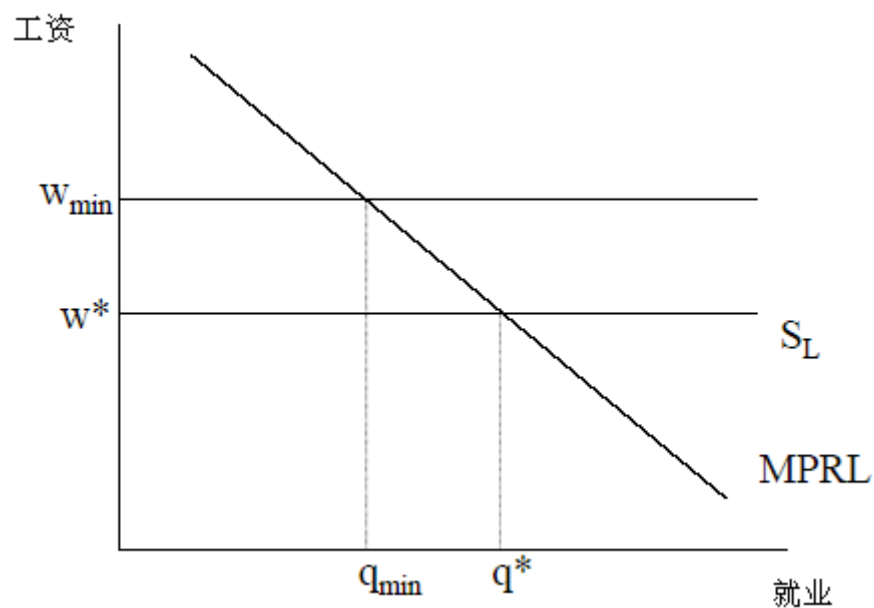
答：需求弹性：

如果工资增长的比例大于（相应的）就业减少的比例 $\Rightarrow$ 总工资上升

$$\eta = \frac{\partial Q}{Q} \frac{w}{\partial w} \geq -1$$

教材中模型的基本假设是什么？

个体“价格接受厂商”假设。



**MPRL** = 劳动的边际收益产品  $\Rightarrow$  “边际工人所生产的量”。它是随雇佣工人增多，而逐渐递减的，这是由于在生产函数中，生产要素的回报是递减的。。

如何确定厂商的原则是：

$$MRPL = w^*$$

回忆厂商的最优问题：

$$\max \pi = pf(L) - w(L)L$$

当  $p$  为产品价格时， $w(L)$  是 “召集”  $L$  工人的必要工资。

$$\begin{aligned} \frac{\partial \pi}{\partial L} &= p \frac{\partial f(L)}{\partial L} - w(L) - \frac{\partial w(L)}{\partial L} L = \\ &= pf'(L) - w(L) - w'(L)L = 0 \end{aligned}$$

这个函数也可以写作：

$$pf'(L) = w(L) + w'(L)L$$

其中：

- $pf'(L)$  = 劳动的边际收益产品(MRPL)
- $w(L)$  = 工资
- $w'(L)L$  = 每增加一单位劳动的工资改变量  $\times$  总劳动力

问：竞争模型中的关键假设是什么？

$$w'(L) = 0 \Leftrightarrow \text{价格接受厂商}$$

当厂商不再是价格接受者时会选择怎样雇佣劳动力？

仍然按照上面的式子：

$$pf'(L) = w(L) + w'(L)L$$

若  $w'(L)L \neq 0$ ，那么厂商必须为它多雇的每一个工人而支付所有工人相同的工资增量。

问：我们怎么知道这意味着更少的就业？

直观上，价格越高，需求数量一定越少。

可选：

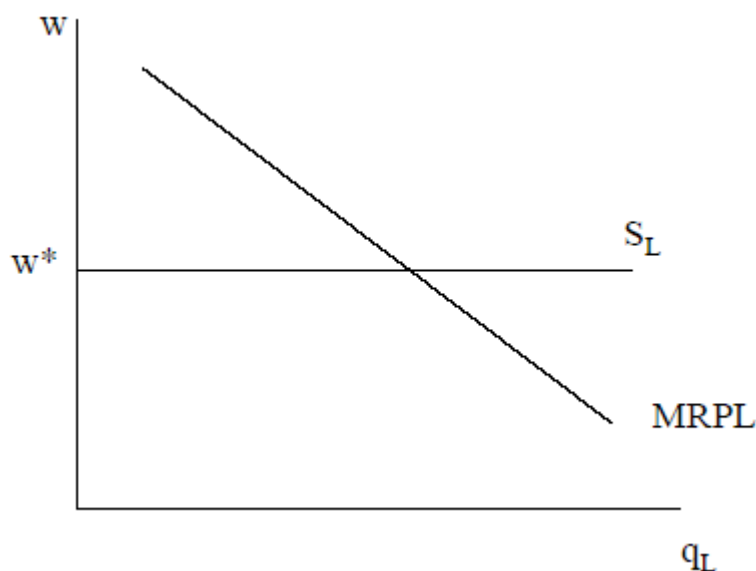
$$\begin{aligned} pf'(L) &= w(L) + w'(L)L \\ w &= MRPL - \frac{\partial w}{\partial L} L \\ 1 &= \frac{MRPL}{w} - \frac{\partial w}{\partial L} \frac{L}{w} \\ 1 &= \frac{MRPL}{w} - \frac{1}{\eta} \end{aligned}$$

其中  $\eta$  是劳动供给的弹性， $\frac{1}{\eta}$  是单个厂商面临的工资关于就业的弹性。在竞争性的劳动市场中，单个厂商面临的工资关于就业的弹性为 0。

$$w = \frac{MRPL}{1 + \frac{1}{\eta}}$$

工资同劳动的边际产品密切相关。劳动供给越有弹性，工资就越低。

## 4.1 传统案例：单个作为价格接受者的厂商



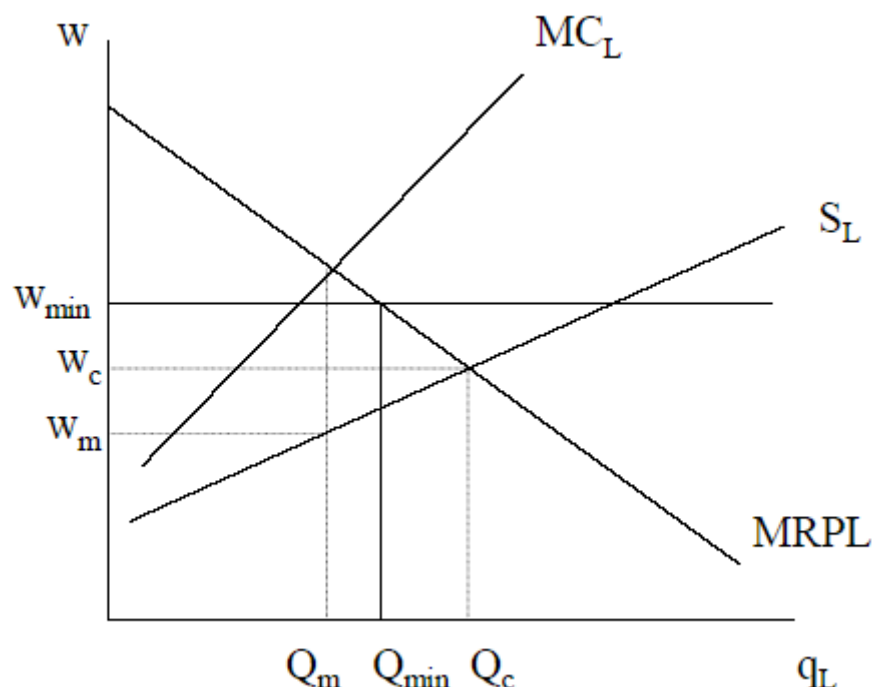
注意劳动供给曲线在市场水平中是向上倾斜的，但对单个厂商而言，它是平坦的。



## 4.2 买方垄断的雇主

买方垄断市场中劳动供给曲线是向上倾斜的。为了增加一个劳动力，垄断买方一定会少量的提高工资。假设所有的工人都接受同样水平的工资（即后进入者不会得到更高的工资），那么雇佣下一个工人的边际成本就不仅仅是他的工资，还包括给其他所有（‘边际内的’）工人所增加的工资。

因此，该厂商的劳动的边际成本曲线就更加的向上倾斜。每个工人增加的成本来自付给该工人的更高水平的工资以及给全部的工人提高的工资。



- 在这个例子中，最低工资标准的实行提高了工资和就业。

$$w_{\min} > w_m$$

$$Q_{\min} > Q_m$$

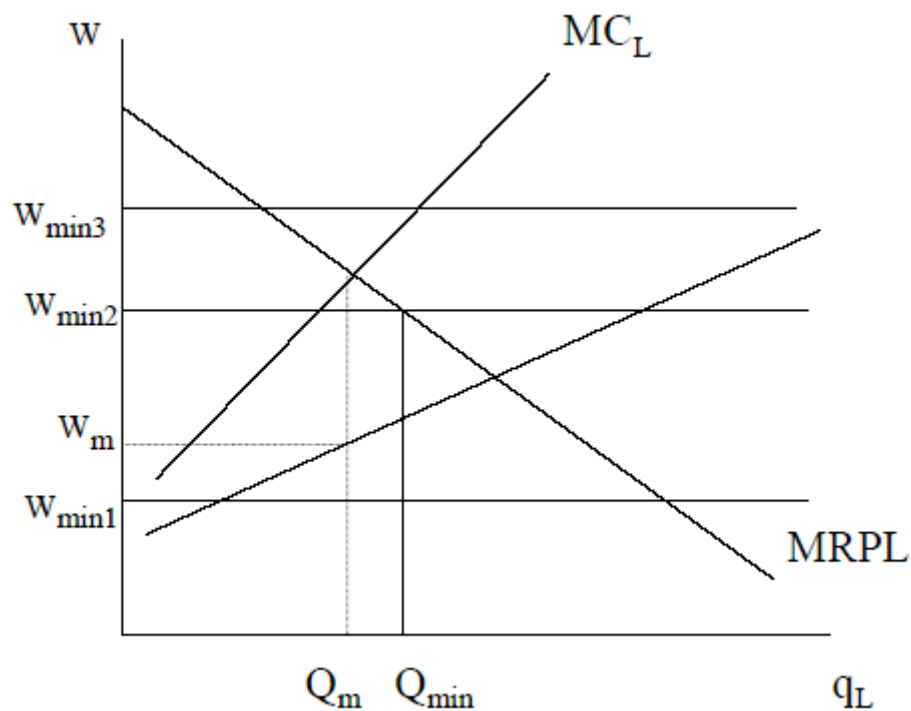
- 原理是怎样的？

此时厂商在  $w^{\min}$  的工资水平上是价格接受者，于是劳动的供给具有“完全弹性”。这就是说，在这个工资水平上有着“无穷多”的劳动可以雇佣（远大于任何一个厂商的需求）。所以买方垄断厂商雇佣劳动的原则是：

$$w = MRPL$$

劳动数量的选择对工资水平没有影响。

问：对买方垄断厂商提高最低工资标准总是能使工资和就业增加吗？



$W^{\min 1}$  - 最低工资低于  $W^m$ ，因此在这个水平上引进最低工资标准没有影响。

$W^{\min 2}$  - 引进最低工资标准增加了工资和就业。

$W^{\min 3}$  - 引进最低工资标准增加了工资但减少了就业。

## 5 买方垄断

若劳动市场中出现买方垄断，你能在哪里发现它？（记住这个原则：厂商自身对劳动的需求改变了市场工资。）

在下面这些市场中我们会看到买方垄断：

- 企业所在地
- 专业化技术人员，例如：IBM 的主机维修工程师
- ‘战俘’ 劳动市场，驻地远离家乡的士兵的配偶
- 新泽西和宾西法尼亚州城镇周围的快餐店？

### 5.1 检验劳动市场中的买方垄断

- 如何检验劳动市场是买方垄断性的还是竞争性的？

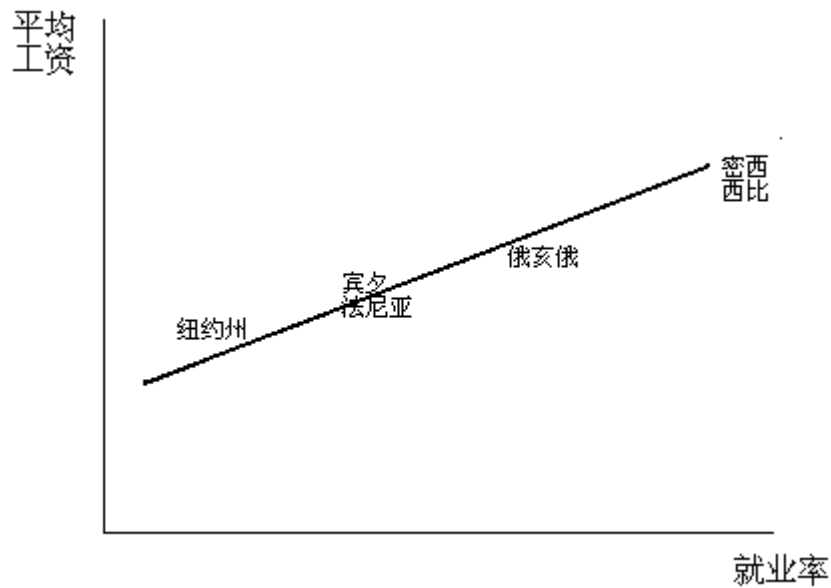
- 问：区别这两个模型最关键的实证含义是什么？

答：在竞争性模型中： $w \uparrow$ （并且） 就业  $\downarrow$

在买方垄断模型中： $w \uparrow$ （并且） 就业  $\uparrow$

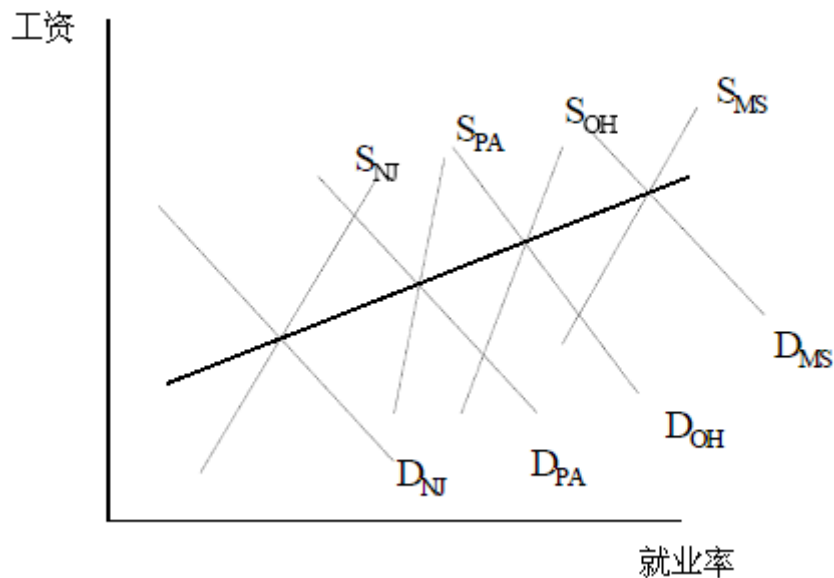
- 那么怎么检验这个特点？

- 我们观察不同的州并且问这个问题：就业提高的地方工资是否也同样提高？  
假设你观察到了如下的情况：



问：这样是不是方便了些？在这个图中工资有什么问题？

答：我们不知道为什么各个州的工资会有所不同，可能是因为需求和供给计划都有着很大的不同。



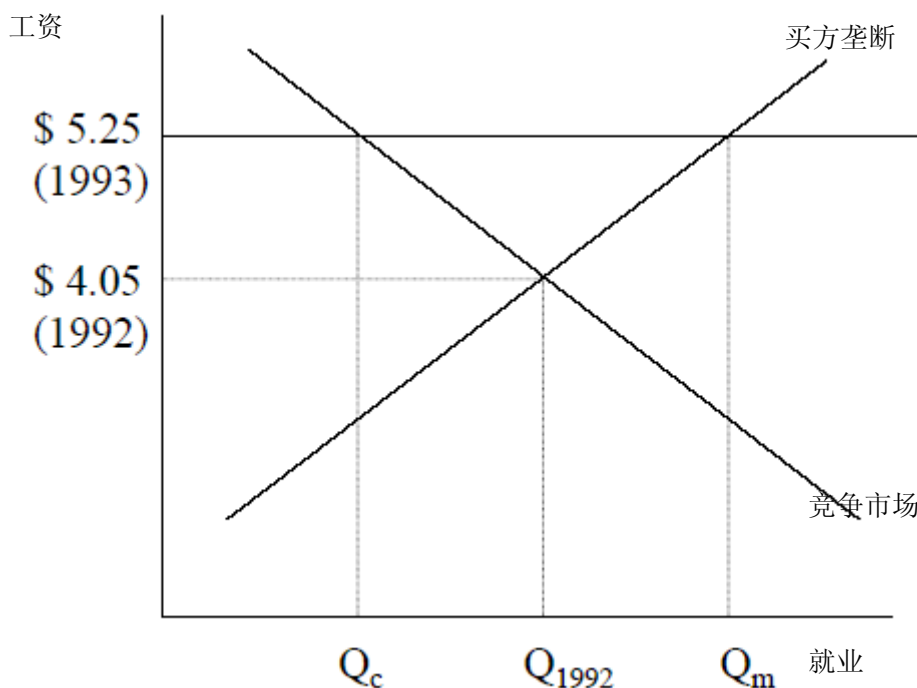
既然就业和工资都是内生变量 - 受供给和需求决定 - 那么这幅图没有告诉我们

任何有关最低工资影响就业的信息。

- 怎样解决这个问题？我们需要做个实验。但是什么样的实验呢？让工资上升。那么让供给增加如何呢？不必。

通过外生变量的调整影响工资进而影响就业，我们能研究工资和就业的关系（向下倾斜 → 竞争市场，向上倾斜 → 买方垄断）

- 一个办法：新泽西最低工资法。



注意在买方垄断市场中通过引入最低工资提高就业的办法只在局部有效 - 意思是，即使是在买方垄断的情况下，最低工资的提高到“过高”也会降低就业。

看看调整最低工资后就业量的变化，我们会发现就业是上升（买方垄断）还是下降（竞争市场）。

注意我们并不看供给和需求曲线。这些只是我们用于启发思维的抽象工具。我们所要观察的是工资水平和就业数量。

在讨论 Card 和 Krueger 的论文之前，我们必须回过头来谈谈社会科学中的因果推理问题。这个问题比乍看起来还要微妙。

## 5.2 社会科学中的因果推理

大部分社会科学（心理学，人类学，社会学，政治学，流行病学和经济学的大部分）重视分析变量之间的相关性 - 如，教育和收入，肥胖和心脏病，幸福和长寿之间的相互关系。

相关性描述的是两个受测变量之间统计数据上的关系。相互关系并不一定等同于原因及其影响。幸福和长寿之间的相关性能精确地测定，然而幸福如何影响寿命却并不为人所知。（或许是同一个的基因控制着长寿和幸福，因此就算更幸福也不会增加寿命。）

在课程 14.03 中，我们一般对这些相关性问题不感兴趣。我们的目标是分析因果联系：（原因：只有通过分析原因和影响的关系而不是仅仅描述和记录相关性，科学才能进步。）

因果联系问题：

- 教育对收入的影响？
- 肥胖对心脏病的影响？
- 幸福对长寿的影响？

这些问题远比相关性问题难以解答。相关性能通过准确的观测数据测量，然而因果联系永远无法直接观察到（有待解释）。

因果联系中有两个因素是相关性研究所缺少的：

- 1 首先，因果联系本质上依赖于时间这个维度。我们说 X 对 Y 有什么影响时，意思是如果给 X 另赋一个值的话，相应的 Y 会取什么值？很容易描述它， $X \in \{0, 1\}$ 。所以，已知  $X = 1$  时  $Y = k$ ，问题是， $X = 0$  时 Y 是多少。为求得一个区别于现实的结果，这个问题的时间回复到了过去。
- 2 其次，因果联系通常是由一些不同的条件所内在衡量。在这个例子中，我们将  $X = 0$  同  $X = 1$  这两种情况进行比较。不区分（某种程度上，或许区分的不是那么明显）我们用来比较的 X 的取值的话，讨论 X 对 Y 的影响是没有意义的；相应地，Y 的取值依赖于 X。

一些注意事项：

设  $Y_i$  为单位 i 的利益，i 可以是一个人，一个细胞，一滴水或是一个主权国家。

我们要考虑两个可能的 i 的结果。如果  $X = 0$ ， $Y_{0i}$  为结果，同样当  $X = 1$  时为  $Y_{1i}$ 。

因此，对每一个单位 i，若对其赋值( $X_i = 1$ )或不赋值( $X_i = 0$ )，我们可以观察到一对典型的结果  $\{Y_{0i}, Y_{1i}\}$ 。

于是单位 i 中 X 对 Y 的影响是  $T_i = Y_{1i} - Y_{0i}$ （T 表示赋值变化的产生效果）。

问题随之而来，即我们无法观察到  $Y_{1i} - Y_{0i}$ 。

取而代之的是，我们考察  $Y_i = Y_{1i} \cdot X_i + Y_{0i} \cdot (1 - X_i)$ 。

这就是因果推理基本问题：

对于同一单位 i，不可能观察到  $Y_{0i}$  和  $Y_{1i}$  的值，进而也不可能直接地观察到单位 i 中 X 对 Y 的影响。

因果推理基本问题的解决办法？

- 1 假定时间稳定性和原因短暂性。如果 X 对 Y 的影响在任何时间点上都是相同的（现在未来）并且 X 对 Y 的影响是不可逆的（所以即使将 Y 对 X 显露也不会改变 X 对 Y 的

影响)，那么我们仅仅通过反复地从 0 到 1 改变 X 的值，就能够观察到  $Y_{1i} - Y_{0i}$ （例如，拨动电灯的开关）。

注意时间稳定性和原因短暂性只是假定；它们不能直接被验证。

- 2 假定单位同类性。如果对于所有的 i，都有完全相同的  $Y_{0i}$  和  $Y_{1i}$ ，那么由于  $i \neq j$ ，我们只用计算  $Y_{1i} - Y_{0j}$  就可以测量因果联系。一般来说，单位同类性不一定是一个可验证的假设。但是在实验室的情况下，作这样的假设是合理的（例如，以两分子水为对象的实验）。
- 3 统计学解决方案。这与其他二者不同。

- 尽管也许永远也观察不到  $T_i = Y_{1i} - Y_{0i}$ ，但我们能够潜在地以某种形式的平均数来估计总数如

$$T = E(Y_{1i} - Y_{0i})$$

- 不过既然我们无法观察到  $Y_{1i} - Y_{0i}$ ， $E(Y_{1i} - Y_{0i})$  的量也不能被直接估测。
- 更加微妙的是：并没有理由去假设  $E(Y_{1i}|X=1) = E(Y_{1i}|X=0)$  以及  $E(Y_{0i}|X=1) = E(Y_{0i}|X=0)$ 。所以，仅仅计算  $E(Y_{1i}|X=1) - E(Y_{0i}|X=0)$  一般并不会得到  $E(Y_{1i} - Y_{0i}) = T$ 。这是因为  $X_i = \{0, 1\}$  成立与否也许同  $Y_{1i}, Y_{0i}$  有关。

更具体地：

$$E(Y_{1i}|X=1) - E(Y_{0i}|X=0) = \underbrace{E(Y_{1i}|X=1) - E(Y_{0i}|X=1)}_{T_i} + \underbrace{\{E(Y_{0i}|X=1) - E(Y_{0i}|X=0)\}}_{Bias}$$

例：比方说 Y 是你上完 4 年大学后在一小时内能认出的数学表达式的数量，X 是个可变的标志，表示你是否进入 MIT 学习。如果我们随机地进行一次数学测试，可以肯定  $\hat{T} = E(Y_1|X=MIT) - E(Y_0|X=NOT MIT) > 0$ ，意即 MIT 的学生在一个小时内比非 MIT 的学生学更多的微积分。

估测进入 MIT 对微积分水平的影响，T 是有效的吗？绝对不是。

擅长微积分的学生选择去 MIT，而不论是否进入 MIT，较一般的学生而言他们微积分本来就学得更好。

所以， $E(Y_{0i}|X=MIT) > E(Y_{0i}|X=NOT MIT)$ 。这就是说，即使这些学生没有进入 MIT，他们也会比那些没有进入 MIT 的学生数学学得更好。因此， $\hat{T} > T$  [见上面的表达式]。这样，即使比较是否进入 MIT 的学生的数学能力，我们也并没有得到一个对于学生进入 MIT 对微积分水平的影响的有效的估测。本质的问题在于进入 MIT 是内生的。学生进入 MIT 部分因为他们擅长数学。所以，假设非 MIT 学生是 MIT 学生的有效对照组是很不合理的。

- 再比方说我们随意地挑选大量的  $i$  并将其中的一半分配给  $X_i = 1$ ，其中的另一半分配给  $X_i = 0$ 。那么这样就可以保证会有  $E(Y_{1i}|X = 1) = E(Y_{1i}|X = 0)$  以及

$E(Y_{0i}|X = 1) = E(Y_{0i}|X = 0)$  进而有：

$$T = E(Y_{1i} - Y_{0i}) = \underbrace{E(Y_{1i}|X = 1) - E(Y_{0i}|X = 1)}_{T_i} + \underbrace{\{E(Y_{0i}|X = 1) - E(Y_{0i}|X = 0)\}}_{bias = 0}$$

在这个例子中，随机化消除了偏向性的因素。由于任意分派，而更有理由相信  $E(Y_{0i}|X = 1) - E(Y_{0i}|X = 0) = 0$ 。

简要地说，通过对  $X_i = \{0, 1\}$  进行处理使它同潜在的结果相独立，随机化过程有可能解决因果推理基本问题： $E(Y_{1i}), E(Y_{0i}) \perp X_i$  意思是  $E(Y_{1i}|X = 1) - E(Y_{1i}|X = 0) = 0, E(Y_{0i}|X = 1) - E(Y_{0i}|X = 0) = 0$ 。这就是“控制组”的原理 - 没有受过处理的组为处理过的组提供了反事实的结果。

为了解决经济学中的因果推理基本问题，我们常常使用统计学解决方案。这是因为，对于人类行为而言，另外两种方法很少有用：时间稳定 + 原因短暂或单位同类。

相反，统计学解决方案几乎总能起作用。然而，出于伦理和/或实际上的原因，我们很少能够运用它……因此，我们寻找一些准实验。

### 倍差法估测

通常我们并不是简单地测量  $Y$  的值，而是把它当作  $X$  的一个函数来测量其变化量。

所以，若我们有一个处理组和一个控制组，我们用表格说明

	前	后	$\Delta$
处理组	$Y_{0i,0}$	$Y_{1i,1}$	$\Delta Y_i$
干预组	$Y_{0j,0}$	$Y_{0j,1}$	$\Delta Y_j$

为什么我们要对比较做出预测？

如果我们有非常大量的处理和控制的单位来进行比较的话，就没有必要做预测。我们只需要简单地计算  $T = E(Y_{1i,1}|T_i = 1) - E(Y_{0j,1}|T_i = 0)$  就行了。

可是一般来说，不会有如此大量的处理和对照个体样本供我们进行比较。

比方说我们正在评估一种新药物对胆固醇水平的影响。每个处理组和控制组选 20 人，给处理组服用该药物，给控制组服用安慰剂，接着比较这两组的平均胆固醇水平。

这个方法没什么问题，不过我们注意到，在偶然的情况下，这两组在实验开始时的胆固醇水平或许会有微小的差别。

在这个例子中，我们也需要获得一些基线数据 - 给药处理之前 - 只为确定它们具有可比性。

若基线的平均是可比的且不相同的。处理组，由于一些偶然性的因素，比控制组的胆固醇水平稍低一点。我们会认为我们的实验倾向于支持药物会降低胆固醇水平（既然处理组得到了一个较好的结果）这个发现。

而正是这种关注使得我们来比较处理组的胆固醇变化和控制组的胆固醇变化。

这种扣除了初始差别之后的方法有可能可以显现出被混同的样本。它能使我们将注意力

集中在处理组相对于控制组的提高上面。

更正式地，比如说未处理：

$$\begin{aligned} Y_{0i,0} &= \alpha_i \\ Y_{0j,0} &= \alpha_j. \end{aligned}$$

我们希望  $\alpha_i = \alpha_j$ ，但是这并不是一定的。

现在，想象经过处理后，我们观察到

$$Y_{1i,1} = \alpha_i + \delta + T_i$$

这里  $T$  表示进行处理带来的影响， $\delta$  表示任何时间都存在的影响。举例来说，胆固醇水平可能会随着人年龄的增长而趋向于恶化。

这样，如果我们设第一种差别为  $Y_{1i}$ ，我们得到：

$$\Delta Y_i = Y_{1i,1} - Y_{0i,0} = (\alpha_i - \alpha_i) + \delta_i + T$$

这个式子显然不能恢复  $T$ ，但是也消除了“初始水平影响”  $\alpha_i$ 。

同样， $\Delta Y_j = (\alpha_j - \alpha_j) + \delta_j$ 。差异化消除了  $j$  组的初始水平影响。

如果我们愿意相信时间的影响对处理组和控制组是等量的话，即  $\delta_i = \delta_j = \delta$ ，那么我们可得

$$\Delta Y_j = Y_{1j,1} - Y_{0j,0} = \delta$$

以及

$$\Delta Y_i - \Delta Y_j = T + \delta - \delta = T.$$

倍差估测法使我们有可能发现因果联系，当处理组和控制组不完全相同并存在潜在的被混同的时间影响时。

### 5.3 现在回到新泽西

令  $Y_{n\alpha, (b,a)}$  为引入最低工资之前和之后的新泽西的就业。

令  $X$  为最低工资的处理。 $X = 0$  为未处理状态（联邦最低工资）， $X = 1$  为处理状态（新泽西最低工资）。

那么，如果我们想要估计最低工资的上涨对新泽西就业的影响的话，我们要计算：

$$\hat{T} = Y_{n1,a} - Y_{n0,b}$$

简单的表示改变前/后的新泽西就业。

这样对因果联系进行估计，我们认为如何？

必要的时间稳定性假设：如果没有最低工资的上涨的话，新泽西就业会保持不变。

这合理吗？很可能不。在我们先前的例子中， $\hat{T} = T + \delta$ 。

那么我们怎么才能改进这个实验呢？

我们可以挑选一些州，他们的状态是任意的，并对其中的一半设置最低工资另一半没有。于是我们能够比较组中每一个州的就业。

这里有个问题是我们无法做这样的一个实验。不过这是个不错的想法。

另一个办法是选择一个我们认为极具可比性的州用做“控制组” - 在本例中，这个州是宾夕法尼亚。



我们在这两个州中取得基线数据，然后将新泽西的变化和宾夕法尼亚的变化进行比较。这就是我们的倍差估测。

### 5.3.1 Card 和 Krueger 的实验

在最低工资如何影响就业水平的研究成果中，这篇论文被广泛引用。

他们的研究结果在经济学家中引起了巨大的争论，而且也导致了 1995 年克林顿政府给数百万工人涨工资，备受争议。

1992 年 4 月 1 日，在新泽西最低工资从每小时 \$4.25 提高到了 \$5.05（这是个相当大的增长）。

而宾夕法尼亚东部（同新泽西接壤）并没有提高最低工资，仍然维持每小时 \$4.25 的联邦最低工资。

Card 和 Krueger 调查了 410 家快餐店。

这个实验的时间段如下：

之前：1992 年 二月 - 三月

之后：1992 年 十一月 - 十二月

设置如下：

	之前	之后	$\Delta$
新泽西	$Y_{n0,b}$	$Y_{n1,a}$	$\Delta Y_n$
宾州	$Y_{p0,b}$	$Y_{p0,a}$	$\Delta Y_p$

$$\hat{T} = \Delta Y_n - \Delta Y_p$$

论文中表 3 反映了“每店的雇员数”

	之前	之后	$\Delta$
新泽西	20.44	21.03	$\Delta Y_n = +0.59$
宾州	23.33	21.37	$\Delta Y_p = -2.16$

$\hat{T} = 0.59 - (-2.16) = 2.76$ ，标准差 1.36（既然  $t\text{-ratio} \approx 2.0$ ，那么 5% 有统计上的意义）。

这篇论文也包括许多别的检验，不过这就是基本的结果：相对于宾夕法尼亚，新泽西获得了  $2.76 \approx 13.5\%$  的就业增长。

解释：

- 1 买方垄断

其他解释：

- 2 饥饿的年轻人
- 3 刺激效应
- 4 被混同的变化（该文并没有考虑对宾夕法尼亚造成的冲击）

## 5 错误的研究对象（为什么选择快餐食品？）