·经/济/科/学/译/丛·

Econometric Analysis of Cross Section and Panel Data

(Second Edition)

横截面与面板数据的计量经济分析

（第二版）下册

杰弗里·M·伍德里奇（Jeffrey M. Wooldridge）著

中国人民大学出版社

·经／济／科／学／译／丛·

Econometric Analysis of Cross Section and Panel Data

(Second Edition)

横截面与面板数据的计量经济分析

（第二版）下册

杰弗里·M·伍德里奇（Jeffrey M. Wooldridge） 著

胡棋智 胡江华 王忠玉 译

中国人民大学出版社

·北京·

图书在版编目（CIP）数据

横截面与面板数据的计量经济分析：第2版/（美）伍德里奇著；胡棋智，胡江华，王忠玉译. 北京：中国人民大学出版社，2016.1
（经济科学译丛）
ISBN 978-7-300-21938-7

Ⅰ.横… Ⅱ.①伍… ②胡… ③胡… ④王… Ⅲ.经济统计—统计数据—计量经济分析 Ⅳ.F224.0

中国版本图书馆CIP数据核字（2015）第226149号

经济科学译丛
横截面与面板数据的计量经济分析（第二版）
杰弗里·M·伍德里奇 著
胡棋智 胡江华 王忠玉 译
Hengjiemian yu Mianban Shuju de Jiliangjingji Fenxi

出版发行	中国人民大学出版社		
社　　址	北京中关村大街31号	邮政编码	100080
电　　话	010-62511242（总编室）		010-62511770（质管部）
	010-82501766（邮购部）		010-62514148（门市部）
	010-62515159（发行公司）		010-62515275（盗版举报）
网　　址	http://www.crup.com.cn		
	http://www.ttrnet.com（人大教研网）		
经　　销	新华书店		
印　　刷	涿州市星河印刷有限公司		
规　　格	185mm×260mm　16开本	版　　次	2016年1月第1版
印　　张	55.75　插页4	印　　次	2016年1月第1次印刷
字　　数	1 164 000	定　　价	128.00元（上、下册）

《经济科学译丛》编辑委员会

目　录

第Ⅳ篇　非线性模型与相关专题

第Ⅳ篇
非线性模型与相关专题

现在，我们把第Ⅲ篇中的一般方法应用到在实际运用中经常产生的一些特定非线性模型研究上。许多非线性计量经济模型的目的是解释受限因变量。粗略地讲，一个**受限因变量**（limited dependent variable）是其取值范围以某些重要的方式受约束的变量。在经济学中所遇到的大部分变量是处于限定的范围之内，但是，并不是所有的情形都需要特殊处理。例如，许多变量——如工资、人口以及食品消费等，仅举这几个例子——仅仅取正值。如果一个严格为正的变量取很多值，那么几乎不需要特殊的计量经济工具。人们常常对变量取对数，然后使用一个线性模型就行了。

当被解释变量 y 是离散的且取有限多个值时，把它处理成一种近似连续的变量就没什么道理。y 的离散性在本质上并不意味着 $\mathrm{E}(y \mid x)$ 的一个线性模型是不合适的。然而，在第 15 章我们将看到，对于二值响应建模而言，一些线性模型具有某些缺点，同时我们将处理一些诸如 probit 与 logit 的非线性模型。我们在第 16 章还涉及一些基本的多项响应模型，包括响应具有一个自然排序的情形。

特别是，当对个体、家庭或者厂商所做的选择建模时，在计量经济分析中会产生其他类型的受限因变量。最优化行为经常导致关于总体的某个不可忽略部分的角点解。例如，在任意给定时间内，处于工作年龄段的人口的一个相当大部分不会在家庭外工作。一年工作小时数在取值范围上具有一个总体分布，但是在零值处有一个堆积。可能当一个线性模型适合于对期望工作小时数建模时，另一个线性模型将有可能导致对一些人预测出负的工作小时数。由于角点解位于零处，所以取自然对数是不行的。在第 17 章，我们将讨论更适合于描述这类受限因变量的计量经济模型。

在第 18 章，我们将涉及计数、分数和其他非负响应变量。我们的重点放在条件均值的估计上，因此聚焦于不需要特定分布假设的估计方法上（虽然在准极大似然分析中它们会名义上设定一个分布）。

在第 19 章，我们转向研究与缺失数据有关的几个问题，包括数据截取、样本选择及损耗。我们可能是首次遇到这样的情形：我们不得不使用一些出现在潜在总体模型中的所有变量的非随机样本。第 20 章处理附加的抽样问题，包括分层抽样与整群抽样。

第 21 章考虑了处理效应的估计，在那里我们明确介绍了对于当代文献而言非常基本的一个反事实设置。这一章揭示了如何将转换回归模型与随机系数模型（含内生解释变量）拟合进处理效应框架。第 22 章以对现代期限分析的介绍作为全书的结尾。

第15章 二值响应模型

15.1 简介

在二值响应模型中，需要解释的变量 y 是一个这样的随机变量，它取值 0 与 1，以此表示某一个事件是否发生。例如，如果一个人被雇佣，那么 $y=1$，否则 $y=0$；如果一个家庭在某个特殊年份致力于慈善事业，那么 $y=1$，否则 $y=0$；如果一家公司拥有一个特殊类型的养老金项目，那么 $y=1$，否则 $y=0$。不管 y 所定义的内容怎样，习惯做法是，当成功时，用 $y=1$ 表示，而当失败时，用 $y=0$ 表示。

如同在线性模型情形中一样，我们经常称 y 为被解释变量、响应变量因变量或者内生变量，而称 $\mathbf{x}\equiv(x_1,x_2,\cdots,x_K)$ 为解释变量向量、回归元、自变量、外生变量或者协变量。

在二值响应模型中，对于 $\mathbf{x}$ 的各种值，关注的内容主要是**响应概率**（response probability）：

$$p(\mathbf{x})\equiv \mathrm{P}(y=1|\mathbf{x})=\mathrm{P}(y=1|x_1,x_2,\cdots,x_K) \tag{15.1}$$

例如，当 y 表示一个就业指示符时，那么 $\mathbf{x}$ 可以包含影响就业状况的因素，诸如教育、年龄、婚姻状况以及其他因素，比如，二值指示符表示是否参与最近的在职培训项目或对过去犯罪行为的测量。

对于一个连续变量 x_j，x_j 对响应概率的偏效应是

$$\frac{\partial \mathrm{P}(y=1|\mathbf{x})}{\partial x_j}=\frac{\partial p(\mathbf{x})}{\partial x_j} \tag{15.2}$$

当用 Δx_j 去乘式（15.2）时，式（15.2）给出了当 x_j 增加 Δx_j，且一旦保持所有其他变量固定不变（对于“很小”的 Δx_j）时，$\mathrm{P}(y=1\mid\mathbf{x})$ 上的近似变动。

当然，比如说，如果对于某些变量 z（例如，z 可以表示工作经验）$x_1 \equiv z$ 且 $x_2 \equiv z^2$，那么我们会对 $\partial p(\mathbf{x})/\partial z$ 感兴趣。

如果 x_K 是一个二值变量，那么关注的内容为

$$p(x_1, x_2, \cdots, x_{K-1}, 1) - p(x_1, x_2, \cdots, x_{K-1}, 0) \tag{15.3}$$

它表示当 $x_K=1$ 与 $x_K=0$ 时响应概率之差。对于我们所考虑的大多数模型来说，无论变量 x_j 是连续的还是离散的，x_j 对 $p(\mathbf{x})$ 的偏效应依赖于 $\mathbf{x}$ 的所有分量。

在研究二值响应模型时，我们需要回想起，有关伯努利（0—1）随机变量的一些基本事实。这里的内容和基础统计学中在设置上的唯一差别是以 $\mathbf{x}$ 为条件的。如果 $\mathrm{P}(y=1 \mid \mathbf{x})=p(\mathbf{x})$，那么 $\mathrm{P}(y=0 \mid \mathbf{x})=1-p(\mathbf{x})$，$\mathrm{E}(y \mid \mathbf{x})=p(\mathbf{x})$，而且 $\mathrm{Var}(y \mid \mathbf{x})=p(\mathbf{x})[1-p(\mathbf{x})]$。

15.2 二值响应的线性概率模型

二值响应 y 的**线性概率模型**（linear probability model，LPM）被设定为

$$\mathrm{P}(y=1|\mathbf{x})=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_K x_K \tag{15.4}$$

像通常那样，x_j 可以是潜在解释变量的函数，这样仅会改变 β_j 的解释。一旦假定 x_1 在函数关系上不与其他解释变量相关，则 $\beta_1=\partial\mathrm{P}(y=1|\mathbf{x})/\partial x_1$。因此，$\beta_1$ 表示给定 x_1 上增加一个单位时成功概率的变化。如果 x_1 是一个二值解释变量，那么 β_1 表示当 $x_1=1$ 与 $x_1=0$ 且其他 x_j 固定不变时成功概率之差。

在自变量中使用像二次项、对数等函数不会引起新的困难。重要的一点是，β_j 现在测量了解释变量 x_j 对特殊概率的效应。

在决定如何选定适当的估计技术时，推导 y 的条件均值与条件方差是十分有用的。由于 y 是一个伯努利随机变量，所以这些内容正是

$$\mathrm{E}(y|\mathbf{x})=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_K x_K \tag{15.5}$$

$$\mathrm{Var}(y|\mathbf{x})=\mathbf{x}\boldsymbol{\beta}(1-\mathbf{x}\boldsymbol{\beta}) \tag{15.6}$$

其中 $\mathbf{x}\boldsymbol{\beta}$ 表示式（15.5）右边的简称。

式（15.5）蕴涵着，给定一个随机样本，y 对 1，x_1，x_2，…，x_K 的 OLS 回归产生了 β_j 的一致甚至是无偏估计量。式（15.6）意味着，异方差是存在的，除非所有斜率系数 β_1，…，β_K 都为 0。处理这一问题的良好方法是使用标准异方差性稳健标准误与 t 统计量。进一步地，还应当利用多重约束的稳健检验。存在着可以使用通常 F 统计量的情形，此时可对所有变量的联合显著性进行检验（同时令常数是无约束的）。由于在这种特殊的零假设下，$\mathrm{Var}(y \mid \mathbf{x})$ 是一个常数，所以此检验是渐近有效的。

如果我们根据 $\mathrm{P}(y=1 \mid \mathbf{x})$ 是方程（15.4）给定的这一假设进行运算，那

么，我们可以通过应用加权最小二乘（WLS）获得一个更有效的渐近估计量。设$\hat{\boldsymbol{\beta}}$表示 OLS 估计量，而且设$\hat{y}_i$表示 OLS 拟合值。于是，倘若对于所有观测值 i，$0<\hat{y}_i<1$，则把所估计的标准差定义为$\hat{\sigma}_i \equiv [\hat{y}_i(1-\hat{y}_i)]^{1/2}$。于是，WLS 估计量 $\boldsymbol{\beta}^*$ 是从

$$y_i/\hat{\sigma}_i \quad 对 \quad 1/\hat{\sigma}_i, x_{i1}/\hat{\sigma}_i, \cdots, x_{iK}/\hat{\sigma}_i, \quad i=1,2,\cdots,N \tag{15.7}$$

的 OLS 回归中得到的。正如在第 12 章对加权最小二乘法的研究中所获得的结果，出自此回归的通常标准误是有效的。此外，所有其他检验都可以利用 F 统计量或者利用加权回归的 LM 统计量来实施。

如果某些 OLS 拟合值没有位于 0 与 1 之间，那么 WLS 分析在没有将偏离的拟合值以特殊的方式调整到单位区间上时是不能奏效的。进而，由于 OLS 拟合值$\hat{y}_i$是条件概率 $\mathrm{P}(y_i=1 \mid \mathbf{x}_i)$ 的估计值，如果所预测的概率是负的或大于 1，那么这就有点难以处理。

除了拟合值位于单位区间之外的问题，LPM 蕴涵着不管 x_j 的初始值如何，x_j 在其他条件不变情况下增加单位长度会改变 $\mathrm{P}(y=1 \mid \mathbf{x})$ 同等的数量。LPM 的这个特征在字面上是不能成立的，因为将 x_j 中的某一个连续地增大最终会导致 $\mathrm{P}(y=1 \mid \mathbf{x})$ 小于 0 或者大于 1。

一个明智的态度是仅仅将 LPM 视为 y 在解释变量上的线性投影。从第 2 章中回忆一下，在 $\mathbf{x}$ 的线性函数中线性投影提供了最好的最小二乘拟合（尽管 $\mathbf{x}$ 本身可能包括潜在解释变量的非线性函数）。如果我们在这个场景中进行运算，LPM 的 OLS 估计是有吸引力的，因为它一致地估计了线性投影中的参数。WLS 估计量可以被视为一个在加权变量上的线性投影，加权变量可能不会多么有意思，特别是因为如果（15.4）不成立时就不能从 $\mathrm{Var}(y \mid \mathbf{x})$ 中得到权重。

如果估计一个二值响应模型的主要意图是通过在 $\mathbf{x}$ 的分布上进行平均，近似估计出解释变量的偏效应，那么 LPM 通常能做得非常好。（关于它表现得有多好的一些证据，可以通过比较 OLS 系数和来自我们在 15.3 节中将转向的非线性模型的平均偏效应得到。）一些预测的概率处于单元区间外的事实不必严重关注。但是不能保证 LPM 提供关于许多不同协变量值的偏效应的良好估计，特别是对于 $\mathbf{x}$ 的极端值。

例 15.1（已婚妇女劳动力参与） 我们使用 MROZ.RAW 中的数据估计已婚妇女劳动参与（$inlf$）的线性概率模型。样本中有 753 名妇女，428 名报告在一年期间具有非零工作小时数。我们用于解释劳动参与的变量是年龄、教育、工作经验、以千元计的非妻子收入（$nwifeinc$）、年龄小于 6 岁的孩子数量（$kidslt6$）以及年龄在 6～18 岁之间的孩子数量（$kidsge6$）；606 名妇女报告没有孩子，而 118 名妇女报告恰好只有一个孩子。通常 OLS 标准误表示在圆括号之中，而异方差性稳健标准误位于方括号之中：

$$\begin{aligned}
\widehat{inlf} = {} & 0.586 - 0.003\,4\ nwifeinc + 0.038\ educ + 0.039\ exper - 0.000\,60\ exper^2 \\
& (0.154)\quad (0.001\,4) \qquad\quad (0.007) \qquad (0.006) \qquad\quad (0.000\,18) \\
& [0.151]\quad [0.001\,5] \qquad\quad [0.007] \qquad [0.006] \qquad\quad [0.000\,19]
\end{aligned}$$

$$
\begin{array}{lll}
-\ 0.016\ age & -0.262\ kidslt6 & +0.013\ kidsge6 \\
(0.002) & (0.034) & (0.013) \\
[0.002] & [0.032] & [0.013] \\
\multicolumn{3}{l}{\ \ N=753, R^2=0.264}
\end{array}
$$

除了 $kidsge6$ 以外，其余所有系数都具有合理的符号，且都在统计上显著，$kidsge6$ 既不在统计上显著，又不在实践上很重要。$nwifeinc$ 上的系数意味着，如果非妻子收入增加 10（10 000 美元），那么成为劳动力的概率预计会下降 0.034。这是很小的影响，因为已知 1975 年收入增加 10 000 美元在此样本中是一个非常大的数目。($nwifeinc$ 的平均值大致是 20 129 美元，其标准差为 11 635 美元。）一旦拥有多于一个小孩子，$inlf=1$ 的概率估计会减少 0.262，这是一个相当大的影响。

753 个拟合概率中的 33 个位于单位区间之外。与其对那 33 个拟合值加以某种调整并应用加权最小二乘法，我们不如使用 OLS 并报告异方差性稳健标准误。有意思的是，这些与通常 OLS 标准误在实践上不重要的方面有差异。

如果 x_j 的大部分均是离散的且仅仅取几个值，对于 LPM 来说的情形甚至会更强。在前面的例题中，为了允许孩子对劳动参与概率的一个减少效应，我们可以把 $kidslt6$ 分解到三个二值指示符中：没有孩子、一个孩子、两个或更多孩子。最后两个指示符可以用来代替 $kidslt6$，以便允许第一个孩子拥有比后面孩子更大的影响。（有意思的是，当使用这一方法时，第一个孩子与第二个孩子的边际效应实际上是相同的。第一个孩子的估计效应大致是 -0.263，同时关于下一个孩子的劳动参与概率额外减少大致是 -0.274。）

在模型饱和的极端情形中——也就是说，$\mathbf{x}$ 包括互斥且穷尽类别的虚拟变量——线性概率模型完全是一般性的。拟合概率正是由 $\mathbf{x}$ 的不同值所定义的每个单元（each cell）内的平均值 y_i；我们并不需要担心拟合概率小于 0 或者大于 1。可参阅习题 15.1。

15.3 二值响应的指标模型：Probit 与 Logit

我们现在研究形式为

$$\mathrm{P}(y=1|\mathbf{x})=G(\mathbf{x}\boldsymbol{\beta})\equiv p(\mathbf{x}) \tag{15.8}$$

的二值响应模型，其中 $\mathbf{x}$ 表示 $1\times K$ 维的向量，$\boldsymbol{\beta}$ 表示 $K\times 1$ 维的向量，同时我们取 $\mathbf{x}$ 的第一个元素为 1。$\mathbf{x}$ 不包括 1 的例子在实践中是很罕见的。对于线性概率模型来说，$G(z)=z$ 是一个恒等函数，它意味着对于所有 $\mathbf{x}$ 与 $\boldsymbol{\beta}$，响应概率不能在 0 与 1 之间。在本节中，我们假定 $G(\cdot)$ 在开的单位区间取值：$0<G(z)<1$，对于所有 $z\in\mathbf{R}$。

式（15.8）的模型通常称为**指标模型**（index model），因为它限定了响应

概率依赖于 $\mathbf{x}$ 的方式：$p(\mathbf{x})$ 仅仅通过**指标**（index）$\mathbf{x}\boldsymbol{\beta}=\beta_1+\beta_2x_2+\cdots+\beta_Kx_K$ 作为 $\mathbf{x}$ 的函数。函数 G 把指标映射到响应概率上。

在大多数应用中，G 是累积分布函数（cdf），有时候它的特定形式可以由潜在的经济模型推导出来。例如，习题 15.2 要求你推导出自基于效用的慈善捐赠的指标模型。如果一个家庭贡献于慈善捐赠，那么二值指示符 y 等于 1，否则为 0。向量 $\mathbf{x}$ 包含了家庭特征、收入以及慈善捐赠的价值（它由边际税率来确定）。在关于特别的不可观测**偏好变量**（taste variable）的正态性假设下，G 是标准正态的 cdf。

G 作为累积分布函数的指标模型可以更一般地由基本潜变量模型* 推导出来，如同例题 13.1：

$$y^*=\mathbf{x}\boldsymbol{\beta}+e,\quad y=1[y^*>0] \tag{15.9}$$

其中 e 表示与 $\mathbf{x}$ 独立的连续分布变量，而且 e 的分布关于 0 是对称的；由第 13 章回想起，$1[\cdot]$ 表示示性函数*。如果 G 是 e 的 cdf，那么由于 e 的概率密度函数（pdf）关于 0 是对称的，所以 $1-G(-z)=G(z)$，对于所有实数 z。因此，

$$\mathrm{P}(y=1|\mathbf{x})=\mathrm{P}(y^*>0|\mathbf{x})=\mathrm{P}(e>-\mathbf{x}\boldsymbol{\beta}|\mathbf{x})=1-G(-\mathbf{x}\boldsymbol{\beta})=G(\mathbf{x}\boldsymbol{\beta})$$

这正好是式（15.8）。

潜变量模型要求 e 服从对称分布并没有特殊的缘由，但是，这恰巧是二值响应模型最经常应用的情形。

在二值响应模型的大多数应用中，最初的目标是解释 x_j 对响应概率 $\mathrm{P}(y=1\mid\mathbf{x})$ 的效应。潜变量公式倾向于给出我们起初对每一个 x_j 关于 y^* 效应感兴趣的印象。正如我们将要看到的，x_j 对 $\mathrm{E}(y^*\mid\mathbf{x})=\mathbf{x}\boldsymbol{\beta}$ 效应的方向与 x_j 对 $\mathrm{E}(y\mid\mathbf{x})=\mathrm{P}(y=1\mid\mathbf{x})=G(\mathbf{x}\boldsymbol{\beta})$ 效应的方向是相同的。但是，潜变量 y^* 极少拥有定义良好的度量单位（例如，y^* 可能用效用单位进行度量）。因此，除了一些特殊情形之外，β_j 的数量大小并不是特别有意义。

概率单位模型（probit model）是满足

$$G(z)\equiv\Phi(z)\equiv\int_{-\infty}^{z}\phi(v)\mathrm{d}v \tag{15.10}$$

的式（15.8）的特殊情形，其中 $\phi(z)$ 表示标准正态密度：

$$\phi(z)=(2\pi)^{-1/2}\exp(-z^2/2) \tag{15.11}$$

当 e 服从一个标准正态分布时，概率单位模型可以由潜变量公式推导出来。

逻辑单位模型（logit model）是满足

$$G(z)=\Lambda(z)\equiv\exp(z)/[1+\exp(z)] \tag{15.12}$$

的式（15.8）的特殊情形。当 e 服从一个标准逻辑斯蒂分布时，此模型可以由

* latent variable model，又称特征变量模型。——译者注

** indicator function，又称标示函数。——译者注

模型（15.9）推导出来。

一般性设定（15.8）允许我们在一个框架下涵盖 probit、logit 以及许多其他二值选择模型。实际上，在下文中，我们甚至不需要 G 是 cdf，但是我们假定 $G(z)$ 对于所有实数 z 严格地位于 0 与 1 之间。

为了成功地应用 probit 与 logit 模型，重要的是要知道，如何既在连续解释变量上又在离散解释变量上解释 β_j。首先，如果 x_j 是连续的，

$$\frac{\partial p(\mathbf{x})}{\partial x_j}=g(\mathbf{x}\boldsymbol{\beta})\beta_j,\quad 其中，g(z)\equiv\frac{\mathrm{d}G}{\mathrm{d}z}(z) \tag{15.13}$$

因此，x_j 对 $p(\mathbf{x})$ 的偏效应通过 $g(\mathbf{x}\boldsymbol{\beta})$ 依赖于 $\mathbf{x}$。如果 $G(\cdot)$ 是严格递增的 cdf，如同 probit 与 logit 情形，那么对于所有 z，$g(z)>0$。因此，效应符号由 β_j 符号给出。同样地，**相对效应**（relative effect）并不依赖于 $\mathbf{x}$：对于连续变量 x_j 与 x_h，偏效应比是一个常数且由相应的系数之比给出：$\frac{\partial p(\mathbf{x})/\partial x_j}{\partial p(\mathbf{x})/\partial x_h}=\beta_j/\beta_h$。在 g 是关于 0 对称的密度的典型情形中，在 0 点有唯一的众数，最大效应正是 $\mathbf{x}\boldsymbol{\beta}=0$ 的时候。例如，在满足 $g(z)=\phi(z)$ 的 probit 情形中，$g(0)=\phi(0)=1/\sqrt{2\pi}\approx0.399$。在 logit 情形中，$g(z)=\exp(z)/[1+\exp(z)]^2$，从而 $g(0)=0.25$。

如果 x_K 是一个二值解释变量，那么由 x_K 在从 0 到 1 且所有其他变量固定不变而引起的偏效应就是

$$G(\beta_1+\beta_2x_2+\cdots+\beta_{K-1}x_{K-1}+\beta_K)-G(\beta_1+\beta_2x_2+\cdots+\beta_{K-1}x_{K-1}) \tag{15.14}$$

这个表达式再次依赖于其他 x_j 的所有另外之值。例如，如果 y 表示一个就业指示符，而 x_j 表示参加在职培训项目的虚拟变量，那么表达式（15.14）表示归功于在职培训项目的就业概率上的变化；这依赖于影响达到受雇条件人的其他特征，诸如教育与工作经验。知道 β_K 的符号就足以确定项目是拥有正的效应还是负的效应。但是，为了求出效应的大小，我们必须估计表达式（15.14）。

对于其他类型的离散变量（例如，孩子数量），我们还可以使用表达式（15.14）的差分。如果 x_K 表示这种变量，那么 x_K 从 c_K 变到 c_{K+1} 的概率效应就是

$$\begin{aligned}&G[\beta_1+\beta_2x_2+\cdots+\beta_{K-1}x_{K-1}+\beta_K(c_K+1)]\\&-G(\beta_1+\beta_2x_2+\cdots+\beta_{K-1}x_{K-1}+\beta_Kc_K)\end{aligned} \tag{15.15}$$

包括一些解释变量的标准函数形式是简单易行的。例如，在模型

$$\mathrm{P}(y=1|\mathbf{z})=G[\beta_0+\beta_1z_1+\beta_2z_1^2+\beta_3\log(z_2)+\beta_4z_3]$$

中，z_1 对 $\mathrm{P}(y=1|\mathbf{z})$ 的偏效应是 $\partial\mathrm{P}(y=1|\mathbf{z})/\partial z_1=g(\mathbf{x}\boldsymbol{\beta})(\beta_1+2\beta_2z_1)$，其中 $\mathbf{x}\boldsymbol{\beta}=\beta_0+\beta_1z_1+\beta_2z_1^2+\beta_3\log(z_2)+\beta_4z_3$。由此可得，如果 z_1 的二次项具有一个驼峰或者 U 形，那么响应概率的转折点是 $|\beta_1/(2\beta_2)|$ [因为 $g(\mathbf{x}\boldsymbol{\beta})>0$]。同样地，$\partial\mathrm{P}(y=1|\mathbf{z})/\partial\log(z_2)=g(\mathbf{x}\boldsymbol{\beta})\beta_3$，因而，$g(\mathbf{x}\boldsymbol{\beta})(\beta_3/100)$ 表示给定 z_2 有 1% 的增长变动时 $\mathrm{P}(y=1|\mathbf{z})$ 的近似变化。含有解释变量之间的交互作用的模