

# AMOS 的使用

- 第一部分:介绍
  - 。 关于文挡
  - 。 访问 AMOS
  - o 文挡
  - 。 获得 AMOS 帮助
- 第二部分: SEM 基础
  - 。 SEM 概述
  - 。 SEM 术语
  - 。 为什么使用 SEM?
- 第三部分: SEM 假设
  - 。 合理的样本量
  - 。 连续和正态内生变量
  - 。 模型识别(识别方程)
  - 。 完整数据或缺失数据的适当处理
  - 。 模型规范和因果关系的理论基础
- 第四部分: 使用 AG 建立和检验模型
  - 。 结构方程——多重回归关系的说明
  - 。 使用 AG 绘制模型
  - 。 将数据读入到 AMOS 中
  - 。 选择 AMOS 分析选项和运行模型
- 第五部分: AMOS 输出解释
  - o 评估整体模型拟合
  - 。 绝对拟合检验
  - 。 相对拟合检验
  - 。 修改模型获得较好的拟合优度
  - 。 浏览路径图
  - 。 独立参数的显著性检验
- 第六部分:摘要:结论的实质性解释

# 第一部分:介绍

# 关于文档

#### SEM=Structural Equation

Amos=Analysis of Moment Structures

本课程使用 AMOS (距结构分析) 软件对结构方程进行简单的介绍和概述。结构方程模型 (SEM) 包括多种统计技术,如路径分析,验证性因子分析,带潜变量的因果关系模型,甚至方差分析和多重线性回归。 课程介绍 SEM 的逻辑,SEM 的假设和输入需求,怎样使用 AMOS 执行 SEM 分析。

到课程结束,<u>能够使用 AMOS 拟合 SEM</u>。也能给出 SEM 适合研究问题的评价和 SEM 方法基本假设的概述。
You must c

应该已经知道使用 SAS, SPSS 或类似统计软件怎样产生多重线性回归分析。也应该理解 怎样解释多重线性回归分析的输出。最后,应该理解基本微软视窗导航操作:打开文件和文件夹,保存文件,重新调用先前保存过的文件,等等。

## 访问 AMOS

可以用下列三种方法访问 AMOS:

- 1. 个人计算机用户须从 SPSS 公司(SPSS 许可版本)或者 Small waters 公司(独立版本)获得许可密码
- 2. 德克萨斯大学的教师,学生和职员经由 STATS 视窗终端服务器访问 AMOS。要使用终端服务器,必须获得 ITS 计算机账号(或分类账号),然后在 NT 服务器上验证账号。接下来下载和配置客户端软件使个人计算机,Macintosh,或 UNIX 工作站能连接终端服务器。最后连接服务器,通过双击位于 STATS 终端服务组中 AMOS程序肖像登陆 AMOS。如何获得 ITS 计算机账号的细节,账号使用的变更,下载客户端软件和配置指导可以在 General FAQ #30: Connecting to published statistical applications on the ITS Windows Terminal Server. 中找到。
- 3. 从 AMOS development website 网站下载免费 AMOS 学生版到个人计算机上。如果模型比较小,免费演示版能充分满足需求。对大型模型,需要购买 AMOS 软件或通过校园网络访问 ITS 共享的软件副本。特别是如果决定利用服务器访问其它程序软件(例如,SAS,SPSS,HLM,Mplus 等等),后一项选择最有效。

#### 文档

AMOS 手册是由 James Arbuckle 和 Werner Wothke 撰写的 *AMOS 4.0 User's Guide*。包含 **20** 多个例子,有许多调查得到的典型模型。这些样本数据的例子在学生版和商业版的 **AMOS** 中都有,所以能容易地在 **AMOS** 中拟合,修改描述模型。

AMOS 4.0 User's Guide 手册在 UT Austin 的职员,学生,教师经检验合格后在 PCL 获得。也可以直接从 Smallwaters Corporation Web site 网站定购。 Barbara Byrne 也写了一本使用 AMOS 的书。 书名是《AMOS 的 SEM: 基本概念,应用和编程》。这本书由 Lawrence Erlbaum Associates, Inc 出版。 Lawrence Erlbaum Associates 有限公司也出版 SEM 杂志(季刊)。杂志包含软件回顾,经验文章,理论文章以及教师部分和书评。

有许多 SEM 的教科书,范围从 Ken Bollen 的百科参考书到 Rick Hoyle 的更多实用编辑手册。几个常用的书名显示如下。

Bollen, K.A. (1989). 潜变量的结构方程 New York: John Wiley and Sons.

Loehlin, J.C. (1997). 潜变量模型 Mahwah, NJ: Lawrence Erlbaum Associates.

Hoyle, R. (1995). SEM: 概念, 问题和应用。Thousand Oaks, CA: Sage Publications.

Hatcher, L. (1996). *循序渐进把 SAS 系统用于因子分析和结构方程模型。*Cary, NC: SAS Institute, Inc.

## 获得 AMOS 帮助

如果在 STATS 视窗终端服务器上访问 AMOS 有困难,请打 ITS 的客服电话 512-475-9400 或发送 e-mail 到 help@its.utexas.edu。

如果能登陆视窗 NT 终端服务器运行 AMOS,但有怎样使用 AMOS 或解释输出的问题,请 预约统计和数学服务部,打 ITS 客服电话 512-475-9400 或发送 e-mail 到 stats@its.utexas.edu。 重要提示: 两个服务部只是在德克萨斯大学的教师,学生和职员使用。有关指导服务的更多信息及常见问题和回答 EFA, CFA/SEM, AMOS, 其它题目见我们的网站 http://www.utexas.edu/its/rc/services。

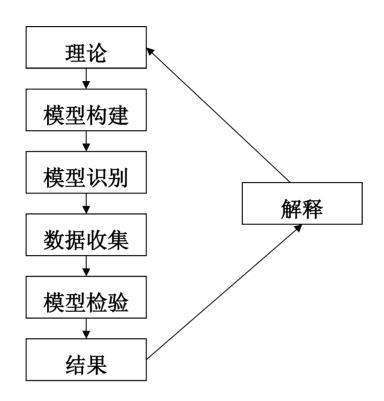
非大学用户和大学 AMOS 用户可以访问 Ed Rigdon's SEM FAQ Web site 网站获得有用的资源;怎样登陆论坛,发表问题和学习更多 SEM 的信息,见 SEMNET 上的在线讨论组

## 第二部分: SEM 基础

## SEM 概述

SEM 是一般线性模型的扩展。它能使研究者同时检验一组回归方程。SEM 软件不但能检验 传统模型,而且也执行更复杂关系和模型的检验,例如,验证性因子分析和时间序列分析。

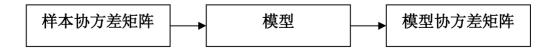
进行 SEM 分析的基本途径显示如下:



研究者首先基于理论定义模型,然后确定怎样测量建构,收集数据,然后输入数据到 **SEM** 软件中。软件拟合指定模型的数据并产生包括整体模型拟合统计量和参数估计的结果。

分析的输入通常是测量变量的协方差阵,例如调查项目得分,虽然有时使用相关阵或协方差和均值阵。实际上,数据分析经常用原始数据提供给 SEM, 程序转换这些数据为它自身使用的协方差和均值。

模型由测量变量间的关系组成。然后,这些关系表示所有可能关系的限制。



结果有模型拟合的全部指数以及参数估计,标准误,模型中自由参数的检验统计量。

#### SEM 术语

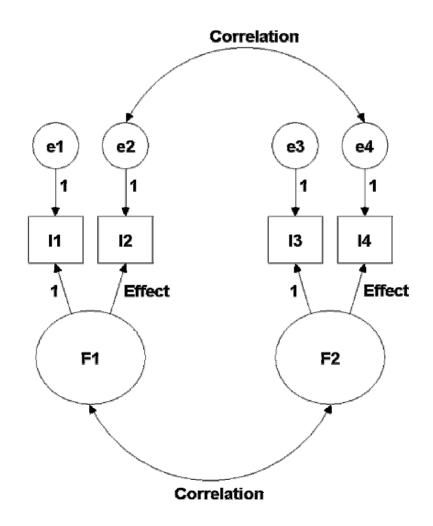
SEM 有其自身的语言。一般来讲统计方法有这种特性,但 SEM 用户和创建者似乎为这种方法发明了一个特殊的语言。

自变量,假设没有测量误差的,称为 ; 因变量或中间变量被称作为 。

直接由研究者测量,而 不能直接测量但可由分析中测量变量的 相关性或关联推断。统计估计可由多种方式实现,探索性因子分析从观测变量的共享方差中 推断潜在因子的存在。

SEM 用户使用路径图表示观测和非观测变量间的关系。椭圆或圆形表示 而长方 形或方形表示 。 残差总是非观测的,所以它们用椭圆或圆形表示。

下图所示,相关系数和协方差由双向箭头表示,表示没有明确定义因果方向的关系。例如 F1 和 F2 是有关的或相关的,但没有声称 F1 导致 F2, 反之亦然。



相反,我们声称 *F1* 引起测量变量 *I1* 和 *I2* 上的观测得分。因果效应在路径图中由单箭头表示。*F1* 和 *F2* 被两个指示因子共享的方差概念化(例如,两个指示因子共有什么。)至今为止正如所猜测的那样, *F1* 和 *F2* 是潜因子; *I1* 到 *I4* 是观测变量。也许它们是调查项。 *E1* 到 *E4* 是 *I1* 到 *I4* 中引起响应方差的残差或误差方差。路径图告诉我们 1 到 4 的得分或调查响应项是由两个相关因子及每项的唯一方差引起的。部分唯一方差或许是由于测量误差。

路径图中一部分路径被标上数字 "1"。意思是这些路径系数设为固定值为 1。必须要包括这些固定值: 它们设置潜在因子和残差的测量尺度。 另一种选择是,把因子方差设置为 1 以获得固有的标准解。注释: 当执行多组分析时不应该使用后一种方法。

## 为什么使用 SEM?

如后面所见,研究者为什么想使用 SEM 和必需用它自己的语言处理一些相当严格的统计假设? SEM 有许多吸引人的优点:

- 假设潜在的统计分析是明确的和可检验的,调查者能全部控制和进一步地分析理解。
- 绘图接口软件创造性地推进和使快速调式模型变得容易(这个特性取决于所选的 SEM 软件)。
- **SEM** 程序同时提供总体模型检验和独立参数估计检验。
- 回归系数,均值和方差同时被比较,即使多个组间交叉。
- 测量和验证性因子分析模型能净化误差,使得潜变量间的关联估计较少地被测量误差污染。
- 拟合非标准模型的能力,包括灵活处理追踪数据,带自相关误差结构的数据库(时间序列分析),和带非正态分布变量和缺失数据的数据库。
- SEM 的最后特征是它最具吸引人的性质。SEM 提供统一的架构,多个线性模型能使用灵活,功能强大的软件拟合。

第三部分: SEM 的假设

#### 合理的样本量

SEM 是一般线性模型灵活有力的扩展。像其它统计方法一样,需要一系列假设。这些假设 应该满足或至少近似地保证有可信赖的结果。

按照 James Stevens 的*社会科学的应用多变量统计*的说法,一个好的经验法则是在标准普通最小二乘多重回归分析中每个因子有 15 个个案。因为 SEM 在某些方面与多重回归紧密相关,SEM 中每个测量变量 15 个个案是合理的。Bentler 和 Chou (1987) 注意到只要

数据表现良好(例如,正态分布,无缺失数据或例外值等等),在 SEM 分析中研究者对每个参数估计,只需要 5 个个案。注意 Bentler 和 Chou 提及每个参数估计要有 5 个个案而不是每个测量变量。测量变量在分析中至少有一个典型地路径系数与其它变量相关联,加上残差项或方差估计,所以认可 Bentler 和 Chou 和 Stevens 推荐的每个测量变量最小要15 个个案相吻合是重要的。更一般的, Loehlin (1992) 使用验证性因子分析模型报告蒙特卡洛仿真的研究结果,参考一些文献后,得出对两到四因子模型,调查者应该收集至少100 个个案,200 更好(如果可能)。使用小样本的结果包括迭代失败(软件不能达到满意的解),不合理的解(包括测量变量的方差估计为负值)和降低参数估计的准确性,尤其是,标准误—SEM 的标准误是在大样本假设下进行计算的。

当数据是非正态分布或在某些方面是有缺陷的情况下\_(几乎总是对个案),需要较大的样本。 当数据有偏斜,有高低峰,不完整或不尽合理时,对所需要的样本量做出绝对的推荐是困难 的。一般的推荐是尽可能获得较多的数据。

## 内生变量的连续和正态分布

SEM 程序假设因变量和中间变量(所谓的内生变量是 SEM 的叫法)是连续分布,有正态分布的残差。事实上,SEM 分析的残差不仅仅要求服从单变量正态分布,它们的联合分布也要服从联合多变量正态分布。然而,这个假设在实际中从未满足。

**SEM** 专家已经开发多种方法处理非正态分布变量。这些方法是为假设有潜在连续分布的变量而设计。例如,也许你管理一个研究参与者自信心项目的李克特量表。量表按照自信的连续程度由低到高计分,即使项目数据不是连续分布,潜在自信分布也是连续的。

相反,其它结果变量不是连续分布。例如,医学研究中病人处理后是生还是死?大部分 SEM 程序目前不能处理这些名义水平因变量的类型。

## 模型识别 (识别方程)

如不久所见,为了产生有判断力的一组结果,SEM 程序需要已知足够数量的相关阵或协方差阵作为输入。另一个要求是方程完全可识别。在 SEM 中,识别涉及参数估计至少有一个唯一解的概念。参数估计只有一个可能解的模型称为 。有无限可能参数估计值的模型叫做 。最后,参数估计多于一个可能解(除了一个最佳或最优解外)的模型叫做 。。

下列方程,来自 Rigdon (1997),或许可以帮助更清晰的理解这些概念:

## x + 2y = 7

在上面方程中,x 和 y 有无穷多个解(例如,x = 5 和 y = 1, x = 3 和 y = 2)。因为 "已知"比"未知"少,所以这些值是欠识别的。恰好识别模型是方程个数与变量个数相同

的方程。

$$x + 2y = 7$$

$$3x - y = 7$$

对这个方程,方程个数与变量个数相同,从而有一对最佳值 (x = 3, y = 2)

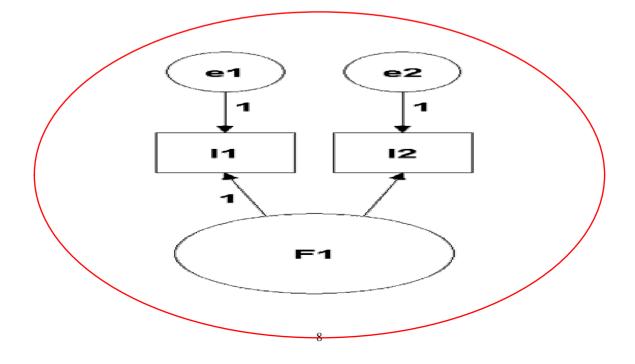
当每个参数可识别,至少一个参数是过度识别时,会出现过度识别模型(例如,方程能由多种方式求解——而不是带一个方程的参数解,多个方程都会产生这种参数估计)。多数使用 SEM 的人都喜欢使用过度识别。过度识别模型有正自由度,也不必拟合恰好识别模型。当有过度模型时,施加在模型上的限制提供一组假设检验,然后它能被绝对拟合模型的卡方统计量和各种描述模型拟合指数估计。与过度识别模型相关的正自由度允许模型有虚假的卡方检验。当过度识别模型拟合良好时,研究者通常认为模型有足够的拟合数据。

为了进行 SEM 分析,识别是结构上或数学上的需要。许多规则能用于评价模型的识别水平,但这些规则不完善,用纸笔计算(事实上,几乎不可能)非常困难,尤其是复杂模型。SEM 软件例如 AMOS 把执行识别检查作为模型拟合过程的一部分。它们通常提供有关欠识别条件的合理警告。

能出现另一个复杂化的情况是经验欠识别。当设置识别模型有非常小 (接近 O)估计的参数估计时,会出现经验欠识别。当 SEM 进行矩阵逆运算时,参数估计在模型参数定义的解空间中下降,这样程序突然侦测到结构欠识别问题是什么。由于 SEM 估计的自然迭代,参数估计例如方差从正值开始在一次一次迭代中渐渐地接近 O。例如,估计值接近 O 的路径系数在 SEM 矩阵逆运算算法中被当作 O 来处理。如果路径系数识别模型是必须的,那么模型就变成欠识别。

各种形式欠识别模式的补救方法是设法找到识别问题的来源和确定来源是经验欠识别还是 结构欠识别。对结构欠识别,唯一的补救方法是重新定义模型。经验欠识别通过收集更多的 数据或重新定义模型来校正。

Rigdon (1997)的例子可以说明这些问题。考虑下列模型:



它包含一个因子 F1,两个误差或残差 e1 和 e2,一个因子载荷连接 F1 到 I2。模型需要估计四个参数:因子的方差,两个误差方差和一个因子载荷。

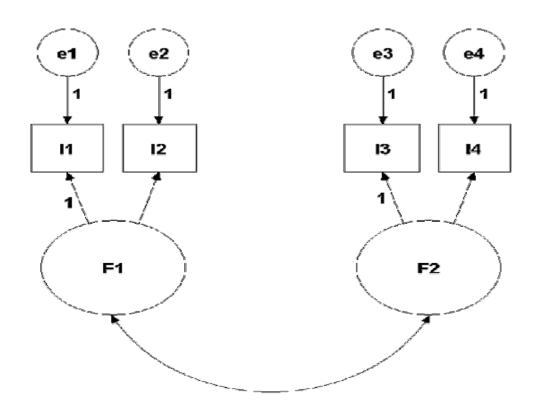
有多少可利用的输入能在分析过程中使用?三个。怎么知道要有输入三个变量?可以使用下面公式

[Q(Q + 1)] / 2

这里 Q 表示模型数据库中测量变量的数目。在这个模型中有两个观测变量, I1 和 I2, 所以由上面公式显示, [2(2+1)]/2=3。有两个方差,一个是两个变量都有的方差,一个是 I1 和 I2 间协方差。

用三个输入估计四个未知参数怎么可能呢? 答案是这是不可能的: 有三个已知变量或可利用的自由度,但有四个未知参数要估计,所以模型的自由度是 3-4=-1 ,明显地不可能。模型是欠识别——为了得到满意识别水平,需要利用这个模型增加约束。

现在考虑第二个模型:



因为模型使用四个观测变量,所以新模型有 [4(4+1)] / 2 = 10 个自由度。从 10 个自由度中减去四个误差方差,两个因子载荷和两个因子方差,一个因子间的协方差结果剩下一个自由度。在结构上模型是可识别的。事实上,因为现在有一个正自由度,所以它是过度识别。

事实表明,如果 *F1* 和 *F2* 的协方差参数估计恰好变成零或非常接近零,模型变成经验欠识别,因为即使在 *F1* 和 *F2* 间通过定义协方差在结构上识别它,从计算机软件观点看,依据经验它不会被识别。

实际上,所有成功的拟合模型是恰好识别或过度识别。通常使用过度识别模型因为这些模型允许检验统计假设,包括整体模型拟合(Loehlin, 1992)。

## 完整数据或缺失数据的适当处理

许多 **SEM** 软件接受相关阵或协方差阵的输入。换句话说,使用另一个软件(例如 **SPSS**), 能自己计算这些矩阵,然后将数据输入到 **AMOS** 或其它 **SEM** 软件中进行分析。例如,在 一本杂志中发表文章,如果打算重新分析报告的协方差阵,这个功能就很有用。

然而,通常使用行数据输入作为首选的模型分析:研究者通过数据库,也许是 SPSS 或其它通用格式(像微软的 Excel),给 SEM 程序,计算协方差作为分析的一部分。如果数据库没有完整的数据,这些程序能够做什么?缺失数据问题的典型解决方法包括个案的列删,如果个案有一个或多个缺失数据,整个个案的记录被删除;对删,两变量相关系数只有用到个案数据时才被计算。对删导致数据库中两变量协方差或相关系数有不同的样本量。缺失数据另一个典型的处理技术是在变量的缺失数据处用变量的均值替代。

但是这些典型的缺失数据处理方法从统计观点来看没有多少吸引力。列删导致统计功效降低,特别是如果许多个案在多个变量上只有几个数据缺失,更不要说在数据库中所有完全测量的个体限制统计推断。对删在边际上比较好,但协方差或相关使用不同的样本量的结果对模型拟合效果有很大的影响,有时包括不可能解。最后,均值替代在所替代的地方会收缩变量的方差,这不是所期望的。这些方法最致命的问题是假设缺失数据是完全随机缺失,实际情况常常不是这回事(Little & Rubin, 1987)。

对缺失数据研究者能做什么?如果缺失数据个案的比例小,比如说小于或等于 5%。列删可以接受 (Roth, 1994)。当然,如果 5%(或更小)的个案不是完全随机缺失,会导致参数估计不一致。另外,缺失数据专家 (例如,Little 和 Rubin, 1987) 推荐在分析中使用极大似然估计方法, 该方法充分利用可使用的数据。在前面的缺失数据中,AMOS 使用极大似然估计。

#### 定义模型和因果关系的理论基础

**SEM** 模型永远不能被接受;它们只能不被拒绝。这导致研究者临时接受一个规定的模型。 **SEM** 研究者承认在多数情况下拟合模型等价于他们自己暂时地接受模型。任何模型可能是"正确"的,因为他们适合数据和首选模型。研究者尽力排除替选模型,扩展替选解释,但这不总是可能的。 因此,使用 **SEM** 需要某些不确定,特别是在控制条件下没有收集的截面数据。(其它通用模型也是如此,例如方差分析和多重回归分析技术。)

因为这个原因,**SEM** 软件需要研究者明确地定义模型。拟合数据较好的模型只能被暂时接受时,拟合数据不好的模型绝对被拒绝。例如,如果用 **10** 个调查项拟合单因子验证性因子分析模型,模型被拒绝,能确信单因子不能充分解释项目的共享方差,一个有用的发现,特别是如果相信一个共有因子没有足够地解释项目的共享方差。假设运行单因子模型,然后在相同的 **10** 个项目上运行双因子模型;前一个模型被拒绝但后一个模型没有被拒绝。现在知道在测量项目中不止一个因素需要解释共享方差。

除了评估单个模型的绝对拟合优度外,也能通过使用比较似然比卡方检验评估竞争模型。回 到前面例子,能使用统计检验相互比较单双因子模型,如果统计检验显著,能得出较复杂的 双因子模型比单因子模型拟合数据更好。另一方面,要是发现两个模型间没有显著差别,能 得出单因子模型与双因子模型拟合数据效果相同。

不是所有模型都可以用这种方法比较。只有在较复杂的模型上利用一组约束或限制得出一个 更简单模型的模型才可以用这种方式比较。这些模型叫做 。 当想比较不能直接比较的模型时,能使用各种描述标准比较这些非嵌套模型。

# 第四部分:使用 AMOS 图形建立和检测模型

SEM——多重回归关系的说明

$$y = i + Xb + e$$

这里 y 是因变量上包含观测得分的向量, i 是表示 y-截距的单位向量, X 是连续分布或分类(编码)自变量的矩阵, B 是回归权重向量, e 表示残差向量或误差或不能由模型解释的剩余得分。

**SEM** 由一系列多重回归方程组成 - 所有方程被同时拟合。事实上,使用 **SEM** 软件能够产生回归分析。

典型的多重回归分析产生几个统计量,包括整体模型的拟合检验和独立参数估计的检验。此外,分析输出非标准回归系数,这些系数的标准误,和标准化回归系数,另外,多重相关系数的平方,或 R<sup>2</sup>表示多重回归方程中自变量解释因变量的方差比例。 正如不久所见,AMOS 产生相同的统计量,虽然是由多个方程而不是一个方程产生的,就像在普通最小二乘回归的情况一样。

(F)

假设有三个连续预测变量的数据库:教育水平,社会经济地位和1967的无力感。有一个连续因变量,1971的无力感。这些数据基于 Wheaton, Muthén, Alwin,和 Summers (1977)大型研究报告的结构所模拟。

如果在 SPSS 中使用这些变量进行多重回归分析,将获得下列结果 (为了节省空间,省略 try it now, but the results is not right 了一些输出):

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.56 <sup>2</sup>	.32	.31	2.62

a. Predictors: (Constant), Powles67, SEI, Educ

#### Correlations

		Powles71	Educ	SEI	Powles67
Pearson Correlation	Powles71	1.00	36	33	.53
	Educ	36	1.00	.54	42
	SEI	33	.54	1.00	32
	Powles67	.53	42	32	1.00
Sig. (1-tailed)	Powles71		.00	.00.	.00.
	Educ	.00		.00	.00.
	SEI	.00	.00		.00.
	Powles67	.00	.00	.00	
N	Powles71	932	932	932	932
	Educ	932	932	932	932
	SEI	932	932	932	932
	Powles67	932	932	932	932

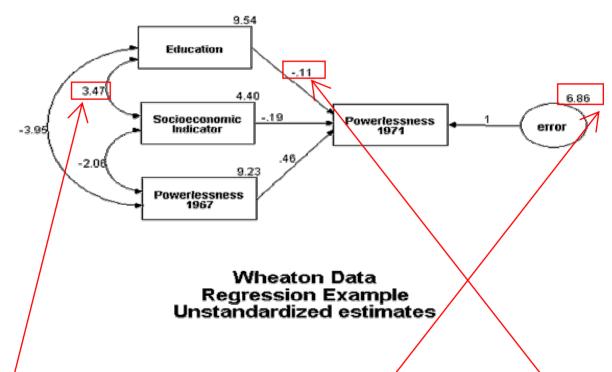
47.00	4747	ien <i>ts</i> a	

		Standardiz ed Unstandardized Coefficient Coefficients s				
Model		В	Std. Error	Beta	t	Sig.
1	(Constant)	6.11	1.02		6.02	.00.
	Educ	11	.03	11	-3.24	.00
	SEI	19	.05	13	-3.98	.00.
	Powles67	.46	.03	.44	14.61	.00.

a. Dependent Variable: Powles71

SPSS 的输出显示每个预测因子和预测因子与因变量的相关系数。 然后显示  $R^2$  为 0.32,在最终表中显示非标准和标准回归系数及显著性检验 。

现在考虑用 AMOS 拟合相同的模型:

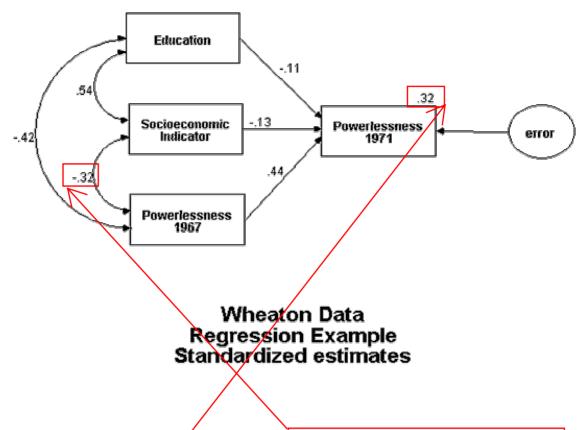


三个预测因子允许共变; 预测因子的协方差显示在图形中。例如, 教育和社会经济地位的 协方差是 3.47。因子的方差出现在长方形的右上方。非标准回归权重,对应着 SPSS 输出中的 *B* 系数,连接 AMOS 路径图中预测变量到因变量的长方形。例如, 教育变量的非标准 回归系数是 -.11。因为非标准回归系数表示预测变量改变一个单位因变量的改变量,这个结果建议教育水平每增加一个单位,从 Wheaton 等人的样本总体中 1971 的无力感下降 -0.11 单位。

路径图也反映 1971 的无力感的刺余方差。因为这个方差不是直接测量的,在 AMOS 图中

用一个潜在变量来表示。误差方差的估计是 6.86。

除了非标准回归系数外,SPSS 也产生标准回归系数和  $R^2$ ,概述收集的预测变量解释因变量的方差比例。AMOS 也产生这些结果,但用单独的图示显示标准系数和  $R^2$  值。



在这个路径图中协方差被相关系数代替。例如,社会经济地位和 **1967** 的无力感的相关系数是-**0.32**,相同的值显示在上面的 **SPSS** 相关系数表中。现在标准回归系数连接预测因子和因变量,因变量的 **R**<sup>2</sup>值出现在长方形的右上方。标准回归系数权重表示预测变量中改变一个单位的标准差引起因变量的改变量。

AMOS 也产生 SPSS 那样的输出表,用于显示非标准和标准回归系数,非标准回归系数的标准误估计,和非标准回归系数等于零的原假设的统计显著性。 AMOS 输出表在下面详细的讨论。

#### 上面显示的回归分析有几方面的限制:

- 不允许有多个因变量或输出变量
- 中间变量不能包含在与预测因子一样的单一模型中
- 预测因子假设为没有测量误差

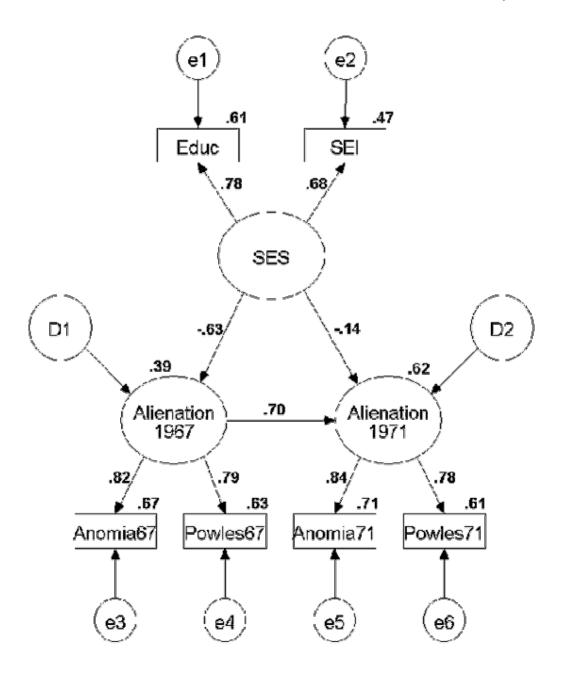
- 误差或残差变量只是模型中允许的潜在变量
- 预测因子间的多重共线性会妨碍结果解释

AMOS 能拟合不受这些限制的模型。Wheaton 等人的模型应用就是一个例子。数据在下面部分给出。

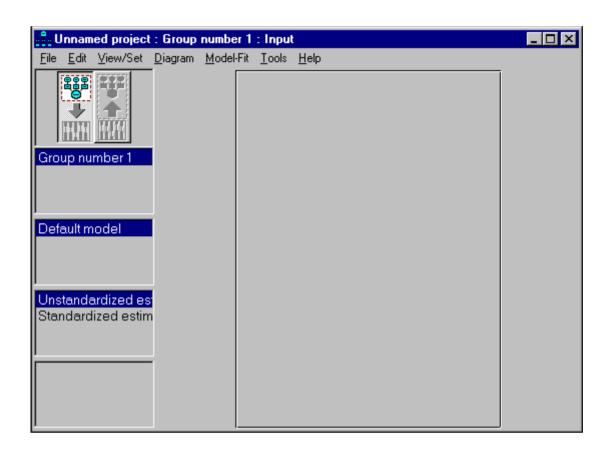
## 使用 AMOS 图形绘制模型

假设想检验随时间改变疏离感稳定性的模型,在 1967 和 1971 年两次测量无价值感和无力感, 及教育水平和社会经济地位。包含标准回归系数的路径图显示如下。





怎样检验这个模型并像在上面显示的那个一样从头开始绘制出版级的路径图?第一步是登录 AMOS。从 Windows 工具栏选择开始一程序—AMOS 4—AMOS Graphics (简称 AG)。现在将看到 AG 的窗口,看上去像这样:



也会看见活动的工具栏。工具栏挡住了 AMOS 一部分绘图区,但是通过拖动能移动它到计算机屏幕的边上。单击鼠标可以选择工具栏。如果鼠标放在工具栏所选的肖像上,AMOS 显示名称和工具的函数。也能右击工具栏肖像获得更多有关如何使用 AMOS 工具的帮助。

要从工具栏上增加或移动一个工具,在 **AG** 的 **Tools** 菜单上选择 **Move Tools**。有时可能需要重新调整工具栏查看所有可用的工具,尤其是如果要显示多个工具或显示器较小时。

用鼠标单击肖像可以激活工具栏。激活或使用的工具会出现被按下或按低的肖像。要取消激活的工具,再次单击它的肖像。肖像恢复到像工具栏上的其它工具一样的状态。

首先,单击激活工具栏选择 Draw Latent Variables and Indicators 工具,然后,单击 Draw Latent Variables and Indicators 按钮使它激活。这个工具类似有三个指示变量的因子分析模型:

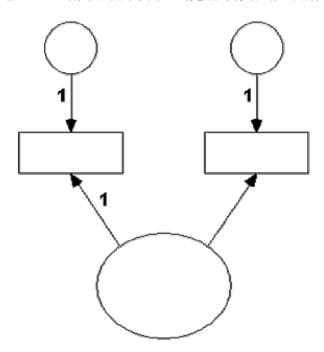


移动鼠标到绘图区,通过按住鼠标键绘制椭圆。一旦对绘制的椭圆感到满意,双击椭圆肖像。出现带两个观测因子的潜在变量。

也许潜在变量和观测变量靠的太近 AMOS 绘图区域的边缘。要向中间移动它们,点击 Preserve Symmetries 按钮。



接下来,点击红色车箱工具 (Move Objects); 然后点击并拖动潜在变量和它的指示因 子靠近 AMOS 绘图区域的中间。当完成时将绘出如下的图示:



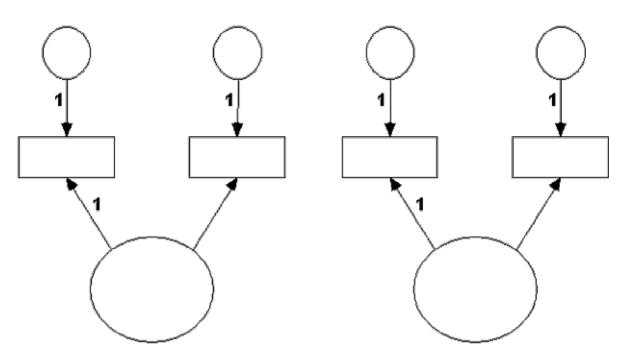
到目前为止,已经绘出了期望模型的一部分。这部分模型能够表示为带两个观测因子, Education,和 Socioeconomic Index (SEI)的潜在变量 Socioeconomic Status (SES)。重新调用表示观测变量或测量变量的矩形,椭圆表示潜在变量或非观测变量。指向 两个观测变量的两个椭圆表示观测变量的残差,而较大的椭圆表示它们的共享方差。共享方 差按照潜在变量 Socioeconomic Status 计算。

模型还有两块类似于这部分(1967年疏离感和1971年疏离感), 所以能复制已经建立的这部分模型。要做这些:

- 点击 Preserve Symmetries 按钮取消它。
- 点击 **Select All Objects** 按钮**(**用伸出四个手指的凸出的手表示**)**。整个 图形的颜色从黑变到篮。

• 点击 Duplicate Objects 按钮 (类型于影印机),点击潜在变量的椭圆肖像, 向右拖动鼠标。出现一个潜在变量——两个指示因子变量结构的复制品。

提示: 当执行这个操作时按住 **shift** 键,图表的复制部分与先前图表部分水平对齐。 当完成这个步骤时,图表显示如下:



点击 Deselect Objects 按钮取消所选的对象。按钮像四个手指缩回的手

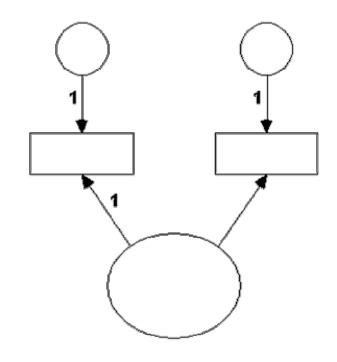


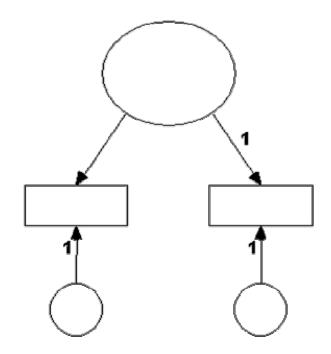
旋转第二个潜在变量的因子以便在 AMOS 图表空间中指向"下"而不是向"上"。选择 Rotate Indicators 按钮。



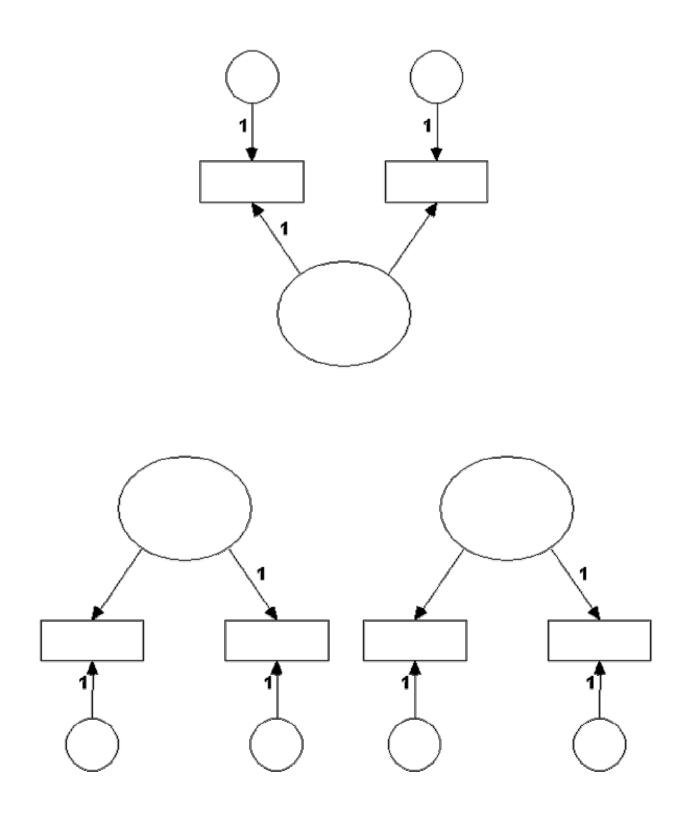
点击第二个潜在变量的椭圆。两个指示因子和它们的残差顺时针旋转 90 度。再次点击该肖像,指示因子再旋转 90 度。点击 Rotate Indicators 工具按钮取消它。

点击 Preserve Symmetries 按钮然后点击 Move Objects 按钮。点击第二个潜在变量并移动它到第一个潜在变量的下方图表的左方。当完成时,图表应该如此:





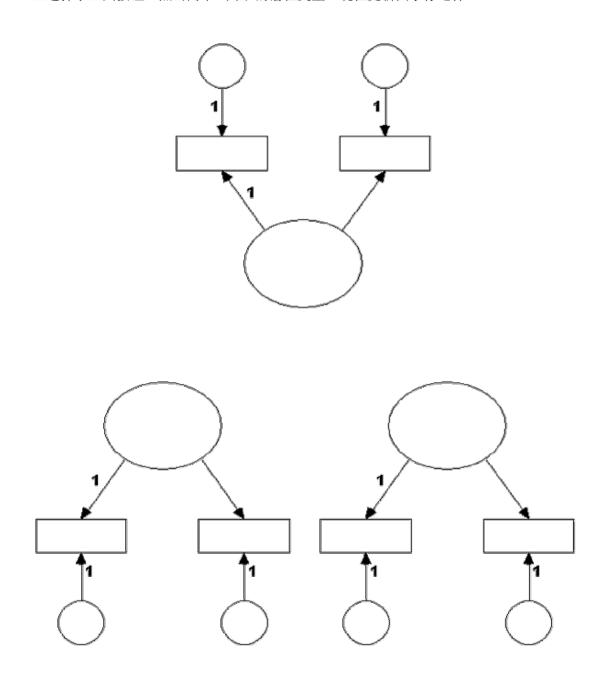
从工具栏再次选择 **Duplicate Objects** 按钮复制下面的潜在变量和观测变量,向右拖动它。现在图表像这样:



注意许多路径有固定值 1。这些值存在保证适当的模型识别。也许希望这些路径值都在变量 集的左边。要做这一点,使用 Reflect Indicators 工具。



一旦选择了工具按钮,点击两个"下面"的潜在变量。现在更新图示像这样:



恭喜你!已经定义模型的测量部分。现在必须告诉 AMOS 在那里找到分析中所使用的数据和必须给观测变量加上标签。

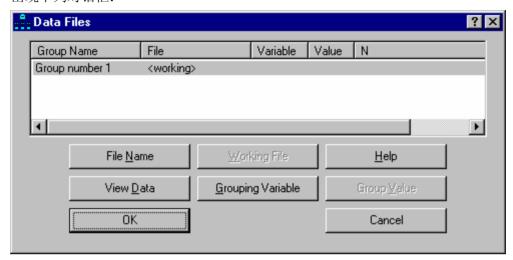
# 将数据读入到 AMOS 中

要将数据读入到 AMOS 中,从 File 菜单选择 Data Files 。

#### File

#### **Data Files**

出现下列对话框:



点击 File Name 定义数据文件名。View Data 按钮浏览外部应用程序的数据文件。例如,如果使用 File Name 按钮定义 SPSS 数据文件,然后单击 View Data, AMOS 将打开 SPSS 以便能看到数据文件的内容。 Grouping Variable 按钮允许在数据库内定义分组变量,能容易地设置和检验涉及多个分组对象的模型。

通常 AMOS 读入下列数据文件格式:

- Access
- dBase 3 5
- Microsft Excel 3, 4, 5, and 97
- FoxPro 2.0, 2.5 and 2.6
- Lotus wk1, wk3, and wk4
- SPSS \*.sav 文件, 版本 7.0.2 到 9.0 (纯数据和矩阵格式)
- 逗号分隔文本文件(在某些国家用分号分隔,逗号用于小数指示符)

AMOS 认可在 Excel 和 Access 中空或空白单元作为缺失数据。系统缺失数据(在 SPSS 数据编辑表中用空单元表示) 也适当地被 AMOS 认为是缺省值。逗号分隔数据文件用两个连续的逗号代表缺失数据。

案例模型的模拟数据是 SPSS 数据文件 Wheaton-generated.sav。如果没有下载这个数据文件,能在下列网站找到:

http://www.utexas.edu/cc/stat/tutorials/amos/wheaton-generated.sav 文件的 Excel 版本来自:

<u>http://www.utexas.edu/cc/stat/tutorials/amos/wheaton-generated.xls</u> 要将这些数据读入 AMOS, 从 AG File 菜单选择 Data Files...。

#### File

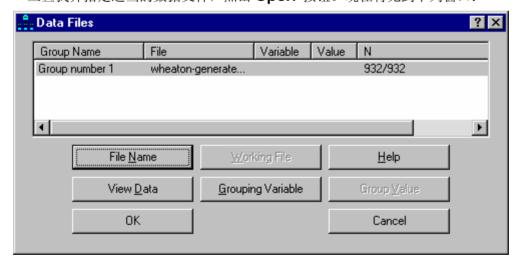
#### **Data Files**

在对话框中,点击 File Name 按钮。 AMOS 产生打开类似下面对话框的文件。



本例中,Wheaton 产生的 SPSS 数据库已经被安装,选择对话框中的 File Name 。可以跟着做,使 AG 复制功能: 在计算机上查找 SPSS 数据文件 Wheaton-generated.sav ,使用 AMOS 选择它。

一旦查找并指定适当的数据文件,点击 Open 按钮。现在将见到下列窗口:

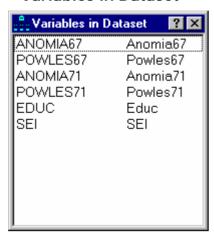


点击 OK 返回到 AG。现在已经得到在 AMOS 中使用的数据。

一旦 AMOS 得到数据,就能给观测变量添加标签。从 View/Set 菜单中选择 Variables in Dataset 选项,窗口显示如下。

#### View/Set

#### Variables in Dataset



每次选择一个变量拖动到路径图长方形区域的适当位置。注意有些变量名太长以致在长方形内放不下。这个问题的一种补救方法是通过点击 Resize Diagram to Fit the Page 工具栏肖像扩大整个模型。



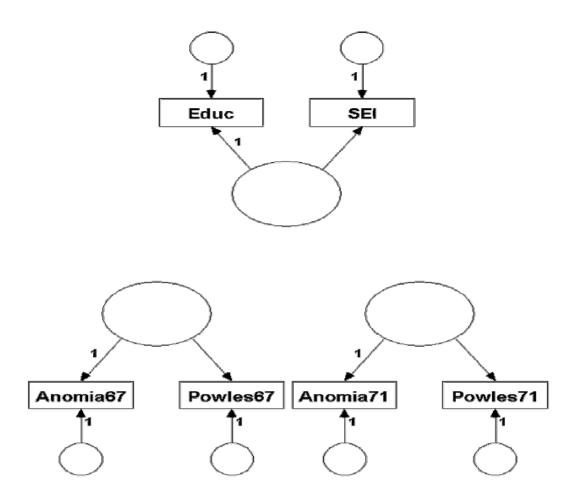
也能使用 Shape Change 工具肖像来改变矩形的形状以便有足够的宽度来容纳变量名:



使用 Select Single Objects 工具选择每个矩形 (它类似一个带伸出一个指头指引的手)。



接下来,点击 Shape Change 工具肖像来改变矩形的形状。最后,通过双击指定变量的矩形能缩小变量名字体的大小。做完这些调整后,测量模型的路径图看起应该来像这样:



#### 还剩下三个问题:

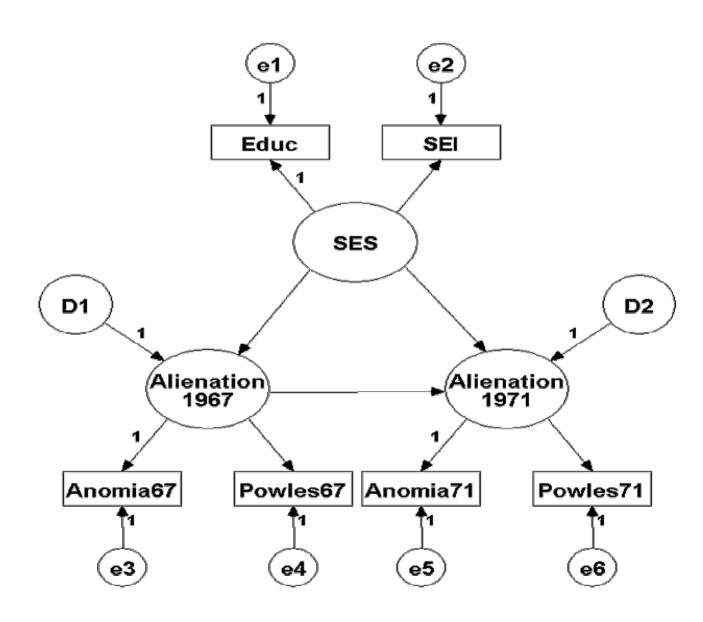
- 1. 潜在变量命名。
- 2. 定义潜变量间的关系(结构模型)。
- 3. 为模型中预测其它变量的潜变量建立残差项。预测另一个变量或一组变量的变量必须有残差。

要命名上面的潜变量,在上面双击它,在 Variable Name 文本框中录入它的名字 (SES) ,然后关闭它。要命名左下方的潜变量,在上面双击它,在 Variable Name 文本框中录入它的名字, *Alienation 1967*,对右下方的潜变量重复这个过程, *Alienation 1971*。 why? i have no this item!

对较大的模型,在 Tools 菜单下有一个 Name Unobserved Variables 宏变量。这个宏自动命名不可观测的变量名。现在选择并运行这个宏来命名没有在路径图中命名的剩余不可观测的变量(例如,残差)。宏运行完后,每个残差将有一个名字,开始的一个是 *e1*。在潜变量间定义路径,使用单箭头表示因果关系,双箭头表示双向关系(相关关系)。要建立潜变量残差,使用 Add Unique Variable 工具肖像。



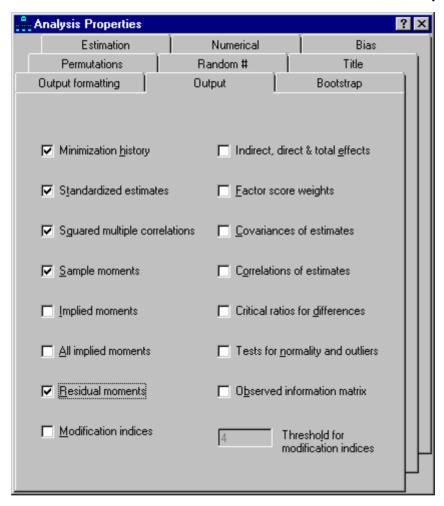
做完这些调整后,完成的模型应该像这样。



绘制模型后,通过选择 File—Save As 保存文件。保存模型时,AMOS 在原始图表的相同目录下自动建立两个备份模型文件。原始图表文件扩展名为 .amw;备份文件的扩展名分别为 .bk1 和 .bk2。

## 选择 AMOS 分析选项运行模型

现在准备运行模型。但是首先,从 View/Set 菜单中选择 Analysis Properties。研究每个选项下可提供的各种各样的选项。最感兴趣的选项之一是 Output 。



在这个分析中,研究者需要许多选项,包括标准估计,复相关系数的平方,样本协方差矩阵和 **AMOS** 拟合模型后剩余残差协方差矩阵。

下一步,检查 Estimation 。注意这个选项提供允许估计平均值和截距的检查框。

如果数据库中的个案有缺失数据,AMOS 需要估计平均值和截距;所以如果数据库中有包括在模型中的观测变量的缺失数据必须选择这个检查框。

因为本模型的数据库不包含缺省数据,目前不用关心平均值,不用考虑 Estimation 设置,以默认值为准。

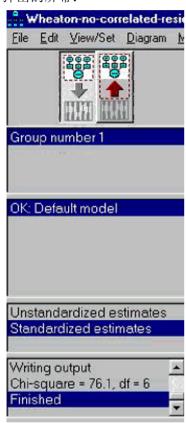
运行模型前,从 File 菜单中选择 Save As 保存文件,在计算机的硬盘上的适当位置保存模型文件的副本。

要运行模型,关闭 Analysis Properties 窗口,点击类似于算盘的 Calculate Estimates 工具肖像:



# 第五部分:解释 AMOS 输出结果

通过检查 AG 绘图区域左边能查看模型运行的是否成功。Wheaton 模型运行后考虑下列 弹出的屏幕:



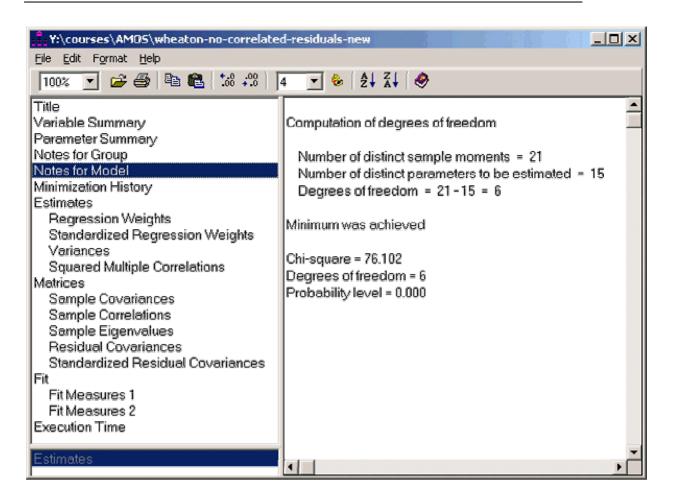
输出的底部显示卡方值为 **76.10** ,自由度为 **6**。检验整体拟合的数据模型的统计检验。这个检验统计量检验数据的总体模型拟合,检验下的原假设是模型拟和数据,所以这个检验希望发现较小的,非显著的卡方值。卡方模型拟合统计量的显著性检验在下一部分描述。

## 评估整体拟合模型

到目前为止,看见的输出结果都不包含统计显著性检验。AMOS产生这样的检验。它们存储在数据表文件中。通过选择 View Table Output 工具能访问这个输出。

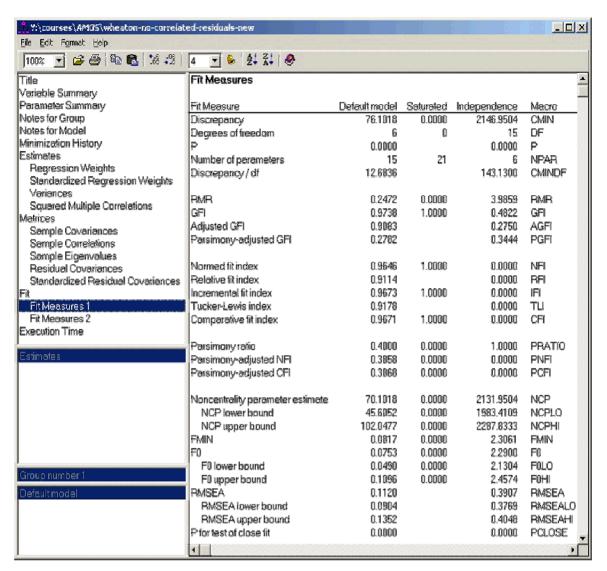


出现的表在窗体的左边包含一栏输出分类,窗体的右边是空白区。当为呈现选择一个特殊的输出子集时,它在右边的空间内出现。考虑输出的 *Notes for Model* 部分。



每次 AMOS 分析完成后,应该总是检查 AMOS 输出的 Notes for Model 部分,因为 AMOS 呈现这部分输出中的大部分错误和告警信息。在上面的输出显示中,AMOS 报告取得最小值,没有错误或告警。报告绝对模型拟合的卡方检验,及自由度和概值。这些结果的解释细节在下面进一步的讨论。

这部分输出结果没有错误或告警意味着进入下一个感兴趣的输出部分,*Fit Measures*,是安全的。





Fit Measures输出包含五列。第一列显示拟合测量的名称。第二列为默认模式,包含AG路 径图中定义的模型拟合统计量。

接下来两列为饱和模型和独立模型,参考两个基线或由 AMOS 作为部分分析自动拟合比较模型。饱和模型包含与可利用的自由度或输入到分析中一样多的参数估计。这样,饱和模型是由 AMOS 拟合限制最少的模型。相反,独立模型是能被拟合的限制最多模型之一:它只包含观测变量方差的估计。换句话说,独立模型假设观测变量间所有关系是零。

#### 绝对拟合检验

整体拟合模型的卡方检验在输出窗口中用 *Discrepancy* 表示。它的值是 **76.102** 带 **6** 个 自由度,如果模型拟合数据的原假设是真实的,偶尔获得大的或较大卡方值,返回的概率小于 **0.001**。

自由度为 6 表示模型过度识别的水平。由前面的公式,有 6 个观测变量,所以有 [6(6+1)]/2 = 21 可利用的自由度。有 6 个残差估计, 3 个因子方差, 3 条路径系数和 3 个因子载荷共

有 **15** 个估计参数。**21** 个自由度**-15** 个估计参数=整体拟合模型的卡方检验的 **6** 个可利用的自由度。

## 相对拟合检验

因为绝对模型拟合的卡方检验对样本量和输入变量分布的非正态性敏感,研究者常常求助于各种描述拟合统计量评估整体拟合模型的数据。在这个框架中,模型可能在一个绝对的基础上被拒绝,然而研究者还是主张一个规定的模型以一定量胜过一些其它基本模型。换句话说,研究者在本文中做的争论是他们选择的模型实际上不如一个基本模型,有代表性是独立模型。然而,与其它模型相比,执行良好的简约模型可能具有实质性的影响。

例如, 塔克-刘易斯指数 (*TLI*) 和比较拟合指数(*CFI*) 把独立模型的绝对拟合与定义模型的绝对拟合进行比较。两个模型的整体拟合差距越大, 这些描述统计量的值越大。

输出窗口有一个单独区域显示过渡节俭调整拟合统计量(parsimony adjusted)。这些拟合统计量类似于多重回归分析中的调整 $\mathbf{R}^2$ : 过渡节俭拟合统计量用很多的估计参数和少量的剩余自由度惩罚大模型。

拟合输出包含一大系列模型拟合统计量。都是设计为检验或描述整体模型拟合。研究者根据他或她偏爱来选择报告的拟合统计量。常用的报告拟合统计量是卡方(Discrepancy),自由度(DF),概值(P),塔克-刘易斯指数(TLI),和近似均方根误差(RMSEA)和置信区间的上下界。通过Tools,Macro菜单也有标准残差均方根(Standardized RMR),但是重要的一点是拟合指数只是比较数据时才可用(它不会显示包含缺失数据的数据)。

这些拟合统计量都有各种经验法则存在。这些经验法则随统计学家新公布的模拟研究而改变。它是各种拟合测量行为的更深一步的文档。

Fit Measures表的最后一栏为宏(Macro),包含 Fit Measures表中由AMOS报告的各种拟合统计量对应的宏变量名。如果想显示拟合指数作为AMOS路径图输出的一部分,这些宏变量可以定义为路径图的一部分。例如,使用CMIN宏变量在路径图中显示拟合模型的卡方。通过从AMOS工具栏中选择Title按钮,宏变量可用于路径图的标题。

Title

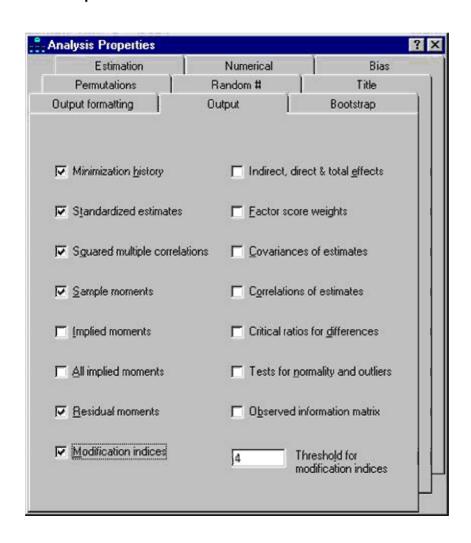
用于产生整体模型拟合的卡方检验的差异拟合函数的公式和描述模型拟合的统计量可以在 AMOS 手册中找到,也可以在 AMOS 程序帮助文件中找到。

## 修正模型获得较好的拟合优度

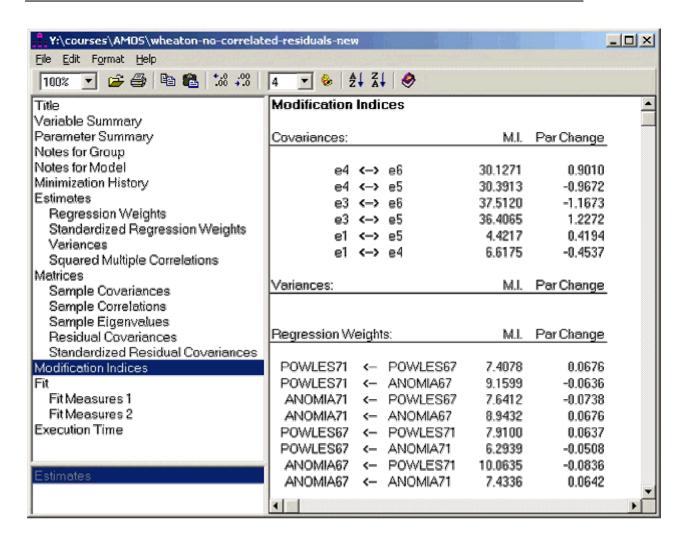
拟合模型与原模型一样是很少见的。有时要获得较好的拟合模型,必须要修改模型。AMOS允许在总体模型拟合中对增加到模型的每个可能的路径使用修正指数产生期望减少的卡方。需要输出修正指数,在Analysis Properties窗口的Output项中选择Modification Indices复选框。

#### View/Set

# Analysis Properties Output



Threshold for Modification Indices (修正指数门槛)允许指定卡方水平的改变需要包含在修正指数输出中的路径。预设值是 4.00 ,因为稍微超过一个自由度卡方分布的临界值 3.84。增加一个参数估计应导致模型卡方至少减少 3.84。修正指数出现在下面。



所有可能的方差都被估计,所以在修正或修订模型中没有被估计的非模型方差。这样,方差部分不包含模型修正信息。然而,合并到重新定义模型中的可能的回归权重和协方差可能导致模型拟合卡方检验统计量的实质改变。

最大修正指数出现在残差协方差的前四对中。例如,如果想用增加协方差办法重新定义模型 然后再拟合模型, *e4* 和 *e6* 的协方差为 0.901。再拟合模型的整体拟合卡方检验应该大约是 30.127 ,小于前一个模型的 76.102。

在修正模型中允许四对误差协方差被估计吗?从修正模型指数结果的来看,答案是"可以的",但在进一步处理前,重新考虑模型修正固有的概念是明智的。理解什么时候基于修正指数的输出修正模型是重要的。您是:

- 1. 重新定义模型
- 2. 基于样本—因变量的结果重新定义模型

任何时候重新定义或修改模型,都会在某种程度上间接地改变它的意义。在许多情况下,界 定模型的改变导致替选模型的真实含义的变更是微不足道或不重要,但在另一些情况下,从 理论观点来看界定模型预示模型的含义必须转换。所以,思考被提议的模型和询问自己改变的模型是否与研究目的一直是非常重要的。

当修改模型时另一个考谅是重点考虑依靠经验数据而不是理论数据帮助定义模型。基于经验修正合并到最终模型越多,模型复制新样本数据越少。为此,应该基于理论及经验结果提出修正指数修改模型。

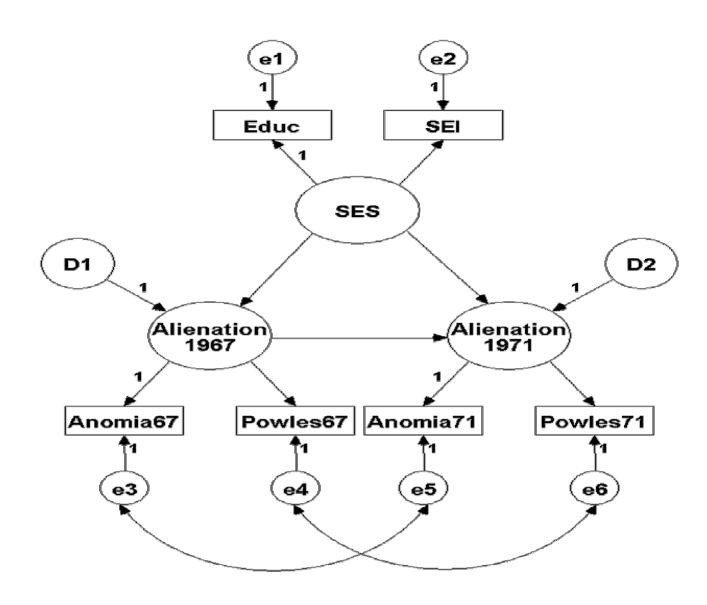
从实际考虑,值得一提的是只有当完整数据输入到程序中时,AMOS 才提供输出的修正指数。换句话说, 当 AMOS 使用缺失数据时不能获得修正指数信息。

能将这些原理用于当前例子的模型吗?可以。检查模型,能看见有两种重复测量的手段:测量 1967 和 1971 年的无价值感和无力感。因为这些数据来自相同的研究参与者,得出 1967 年无价值感和 1971 年无价值感有共享方差似乎是可信的,它不是由当前模型获取的。类似地,也可以共享 1967 年无力感 和 1971 年无力感方差,它也不是当前模型考虑的。能关联这两组变量的残差合并这些共享方差到模型中。要关联残差,返回AG窗口选择由双箭头表示的Draw Covariances工具。

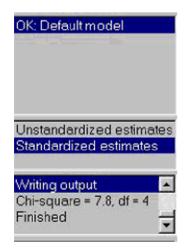


在 1967 年无价值感的误差项 *e3* 和 1971 年无价值感的误差项 *e5* 间拖动相关系数。对无力感残差重复这个过程。

*提示: 通过使用形状改变工具改变相关线条中曲线的大小* 修正模型显示如下。

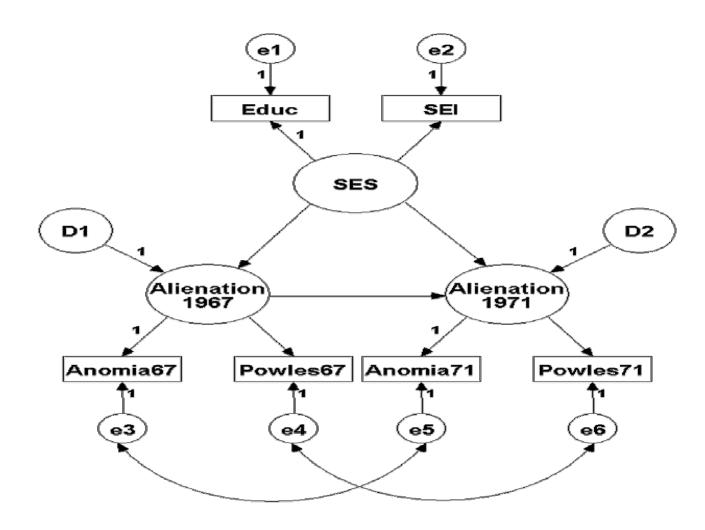


当模型运行时 AMOS 显示这个结果:



包含两个相关残差导致模型拟合卡方从 76.1 到 7.8 实质性地下降。注意卡方检验的自由 度从 6 个自由度降为 4 个自由度。有两个自由度的改变,因为模型的每个参数估计消费了一个自由度;因为在修正模型中包含两个新参数 (两个残差相关系数),新模型保留两个较少的自由度。

由 AMOS 产生的模型拟合表现是如下:



4 个自由度的卡方 7.817 在 0.05 水平下是不显著的: 它的 *p*-值是 0.099。 这一发现表明模型拟合接受 Wheaton 等提取的总体样本中数据。确凿的证据由 RMSEA 拟合统计量提供——所获得的 0.032 明显地低于分界点 0.06。类似地,0.993 的塔克-刘易斯指数是大大高于表示模型拟合满意的极限值 0.95。

一旦获得拟合良好与理论上一致的模型,就可以解释参数估计和每个参数估计的独立显著性 检验。AMOS 提供两种方法检验参数估计。一种方法使用路径图输出可视化参数估计,而 另一个途径使用类似包含总体模型拟合统计量的表格。

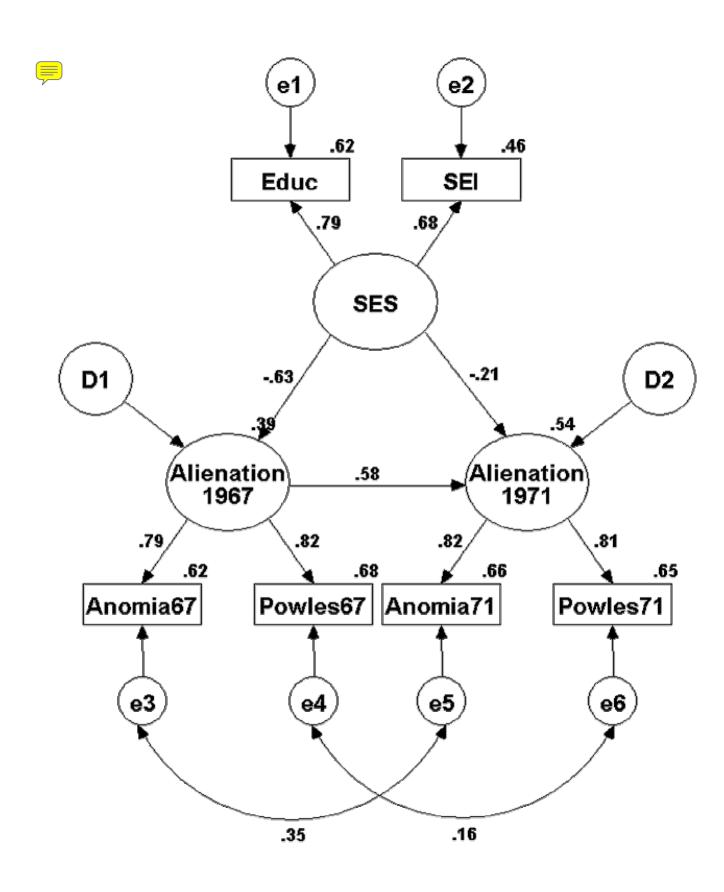
#### 浏览路径图输出

在 AG 窗口的顶端有个紧挨着向下箭头的向上箭头。



点击向上箭头使 AMOS 显示参数估计。 非标准化和标准化估计可通过点击适当的选项选择。标准估计显示在下面的图中。

点击向下箭头返回 AG 画面,在这里能修改现存模型然后再次运行它,或能打开一个新模型或预存的模型文件。当点击向下箭头时,下列参数显示为输出的一部分。

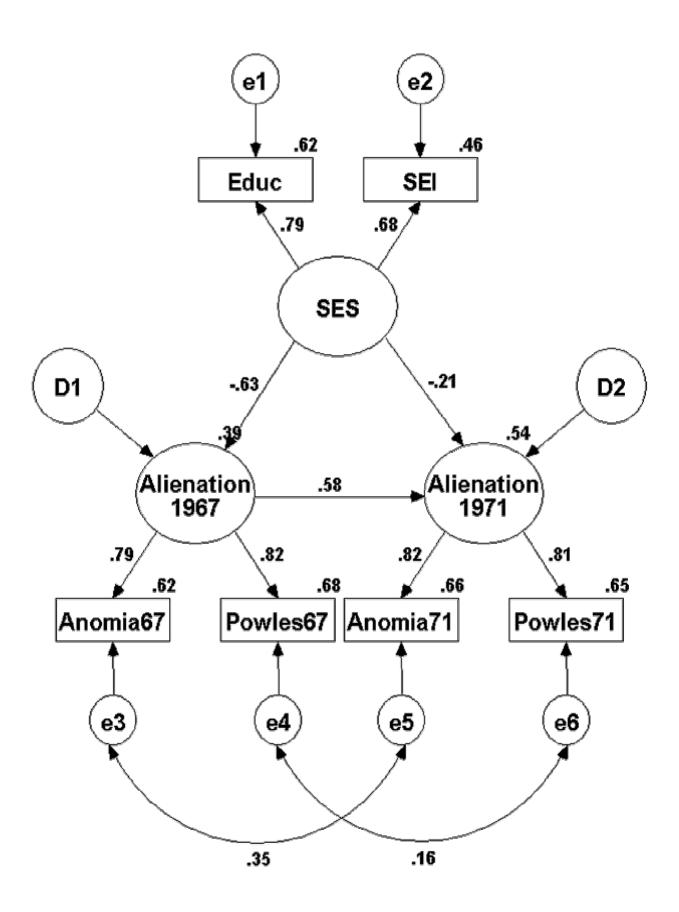


与每个路径相关联的值是标准化的回归系数。这些值表示 X 改变一个标准差时 Y 的改变量。 (对应的非标准化系数表示 X 改变一个单位时 Y 的改变量 )。

在变量上部,AMOS也输出表示因变量或中间变量的R<sup>2</sup>。例如,Anomia67的是0.62。虽然在预设显示中AMOS非常有序地将系数布置好,但移动特殊的参数估计以便出现较少的堆积和更容易解释有时也是必须的。在输出图中要移动参数,使用**Move Parameter**工具。



在不需要的变量上选择工具和移动鼠标直到变量出现红色。然后在希望的地方点击并拖动鼠标以便参数显示的更适当。在现在的路径图中一个较好的选择是潜变量 *Alienation 1967* 的 R<sup>2</sup> 为 0 .39 。通常,它被 SES 到 *Alienation 1967* 的路径覆盖。通过路径对象的右边稍微拖动参数估计对象, 能看见 0.39 没有被阻挡。

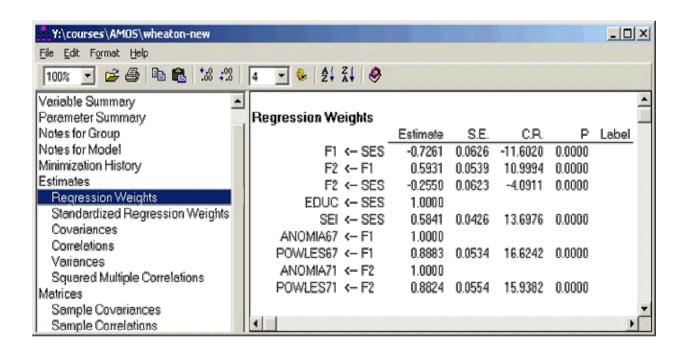


AMOS 的一个优良特征是它的高质量的绘图输出。复制它到剪贴板上,能得到这个输出。 然后把它插入到字处理器中,例如微软 Word 或 PowerPoint。

模型有几个感兴趣的特征值得注意。首先包含潜在(非观测)变量和显在(观测)变量。其次,包括用单项箭头表示潜在变量间的因果关系,和相关系数或几个残差间双向关系。双向箭头分别表示连接 e3 , e5 和 e4 , e6。如上讨论,因为两个无价值感和无力感测量是一样的,是相同的研究参与者不同时间的测量。残差间相关系数说明附加的共享方差。

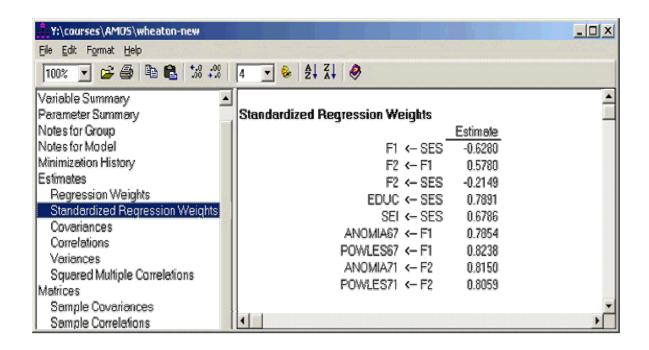
## 独立参数的显著性检验

AMOS 输出也显示非标准和标准回归系数。非标准系数和相关检验统计量出现在下面。非标准回归系数表示自变量改变一个单位因变量或中间变量的改变量。例如,图中数据显示在 SES 中每增加 1 个单位, 1967 年的疏离感下降-0.726。



这张表显示非标准估计,标准误 (简写为 S.E.),和估计除以标准误 (临界比率,简写为 C.R.)。与检验是零的原假设关联的概值显示在 P 列。模型中的所有回归系数在 0.01 水平下与零有显著性差别。

标准估计允许评估结果变量对预测变量的相对贡献。拟合模型的标准估计显示如下。



在这个例子中标准和非标准系数没有太大的差别,可能是因为单位是由调查测量项而来。相反,录入到相同模型不同测量尺度的变量输出的标准和非标准回归系数有明显的差别。

# 第六部分: 摘要:结论的实质性解释

一旦获得拟合良好的模型,与理论一致并提供统计意义显著的参数估计,必须按照研究问题 对它进行解释,然后将结果摘录出来发布。

模型显著的特征包括 1967 和 1971 年的 SES 和疏远感的负相关系数,通过统计意义显著的非标准回归系数进行说明。标准系数揭示 1967 和 1971 年 SES 和疏远感有较强的相关。如所期待的一样,1967 和 1971 的疏远感的相关性也很强。模型的测量部分也是相当好一一最小  $\mathbf{R}^2$  是 0.46 (*SEI*),对行为科学研究,这是一个合理值,其它  $\mathbf{R}^2$  比较高,表示模型在测量项中占较大的方差比例。

值得注意的是即使这个模型拟合数据良好,提供理论一致性也是重要的,可能有其他相等的模型同样好地拟合数据。也许也有拟合数据比这个模型还好的非等价的替代模型。只要可能研究者应该严格检验和排除替代模型。

Hoyle , Panter (in Hoyle, 1995) 和 Hatcher (1994)提供怎样向杂志和教科书编写 SEM 结果的精彩讨论。 Hoyle 和 Panter 提供 SEM 的编写指导,而 Hatcher 给出提交给出版的学术性杂志的手稿的 SEM 分析结果的模拟范文。教科书的全部引用文档显示在后面的参考部分,通过 UT 图书馆系统可以获得。

## 在本课程中学到

- SEM 是什么? 在哪里找到 SEM 资源。
- SEM 术语。
- SEM 的数学基础和它的假设。
- 怎么使用 AMOS 绘制模型。
- 怎么运行 AMOS 模型,评估几个 AG 和文本输出的关键成份,包括整体模型拟合和 检验独立路径系数的统计量。
- 怎么修正和重新定义一个非拟合模型。