
欧盟影子经济的测度及其意义

Peter von der Lippe¹

内容摘要：本文在讨论“影子经济”定义的基础上，将它与“黑色经济”进行比较，与此同时，对如何估测影子经济的各种方法进行了对比分析，最后，对估测影子经济的实证结果进行了评价。

关键词： 欧盟 影子经济 测度

2009年4月在武汉大学举行的“中德人力资源、劳动力市场和社会保险专题研讨会”上，我的演讲论述了以下3个主题：

1. 对影子经济（shadow economy, 简称为SE）的界定，尤其是作为影子经济一部分的地下活动或“黑色”活动，并涉及有关影子经济特别是“黑色经济”（black economy, 简称为BE）的原因与结果的理论；

2. 对影子经济规模的估测方法，并且侧重于对“黑色经济”²的测度；

3. 这些估测的经验结果以及有关这类估测合理性的某些争议。

接下来，把重点放在第二个题目上，并试图评论最频繁使用和讨论最多的测度影子经济的净产值（增加值）的方法。用影子经济占官方的国内生产总值（GDP）的百分比来表达这个所谓的影子经济的“规模”，这种做法相当常见。由于影子经济的产出（SE-production）只是部分地而非全部³地被排除在官方的GDP之外，因此，这种方法只有被描述性意图⁴证明是合理的。

在本文的第二部分中将表明，度量影子经济的“技术状况”（the state of art）几乎并不令人满意，由此带来了长期的有关方法论的争议。重点将被特别放在所谓的货币方法论上。在第3部分，我们对不同国家和经过一段时间过程的影子经济的规模所做的统计估算作一个评论。鉴于所公布的对70多个国家的大量数据的估算是建立在不同的方法之上，我们也提出一个论点对所报告的估算的合理性进行讨论。

一、涉及“影子经济”的定义和理论的一些说法

1. 术语问题

以Buehn et al. (2009)为依据，影子经济（SE）“包括以市场为基础的、合法生产出来

¹ Peter von der Lippe, 德国杜伊斯堡-爱森大学默卡托管理学院教授, 电子邮箱: peter@von-der-lippe.org

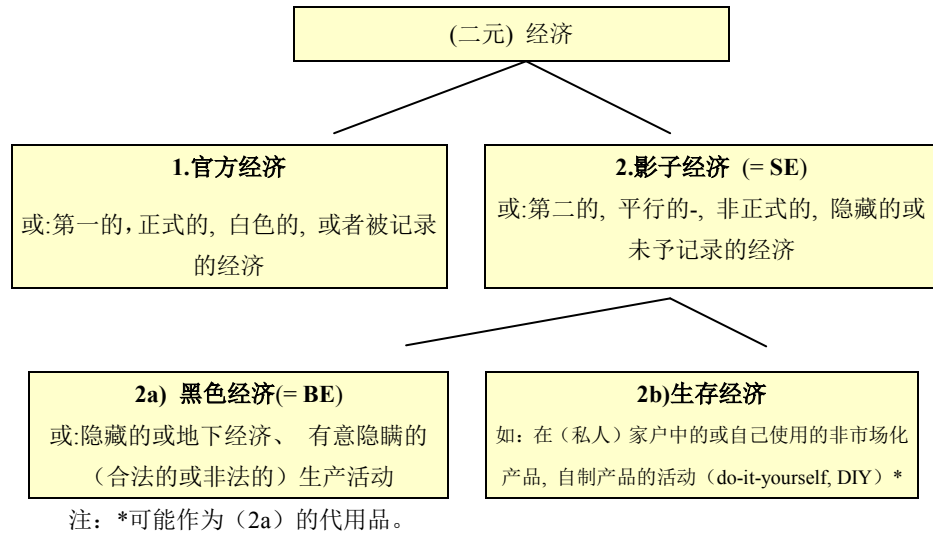
² 那意味着或多或少忽略了所有（未被支付报酬的）非市场产出，这部分产出作为影子经济的另一部分由私人农户自己使用。

³ 被排除的仅仅是图1中计划的2b部分，基本上包括私人农户中的生产性活动。这些东西被正确地排除，因为它们被明确地认为是处于国民账户的“生产边界”（production frontier）之外。由于官方的GDP把2a部分（BE）包括了进去，把SE定义成所有没有记录在GDP中的产出活动是不正确的，尽管这种做法并不是不常见。

⁴ 按照GDP来表示SE的规模在表面上“更好理解”，但是其隐含的解释是：所估计的SE表明，通过哪一个百分比，GDP中的比重应该更高一些，如果它是被（所谓地）正确度量的话。然而，这一解释的错误一目了然，其理由很简单，即在SE中，对BE部分的估计被包含在官方的GDP中（不管它们具有合法的还是非法的特性）。然而，作为一种规则，这类有关BE的数字没有单独地被公布出来（因为它们与GDP不同，它们不符合官方统计准确性的最低要求）。

的产品和服务，这些产品和劳务被有意隐瞒政府当局以逃避支付所得税、增值税或其他税收以及社会保险缴纳金；进而绕开了某些劳动力市场标准，例如最低工资，最高劳动时间，安全标准，等等”。虽然这个定义囊括了典型的影子经济活动，但它并不令人满意，因为部分影子经济显然是不“合法”而且并非“基于市场”。在图 1 中引入了一些必要的区分。

图 1 第二类经济活动的分类



正如几乎在每一个经济中，第二经济或影子经济（参见图 1 中的 2）可以被视为与第一经济共存，这如同一些作者所说的“二元经济”（dual economy）。然而，这一术语远远没有被广泛接受或持续地得以应用。“影子”经济这个词（SE，参见图 1 中的 2）⁵因其臭名昭彰被认为应该更适合称之为地下经济或者“黑色”经济（BE，即图 2 中的 2a）。诸多混淆应归结于对于合法和非法产品及其生产之间的区分不的明确。

“月光下的活动”（moonlighting）⁶（晚上或周末向他人提供付酬的服务，然而却不以做第二份工作的形式向税务当局申报）也许构成“黑色”生产的最显著的案例。这里，并非劳务本身（例如：修理电视机或汽车）不合法，而是逃税使这种服务变成了“黑色”的。其他同样被称为“黑色的”活动⁷的这类非法的（或至少是不道德的）活动，例如非法贸易（运输毒品，黑市购酒、麻醉品（narcotics）等），走私、卖淫、奴役他人、贿赂、欺诈、收受赃物⁸、伪造文件或贪污（滥用公共权力以谋求私利）。然而，“黑色”经济活动并不必然是违法的或者甚至是犯罪的行为。

另外，影子经济（或 BE）突出的一些案例是非法（对外来者）雇佣（常与非法移民以及旨在避税、逃避社会保障缴纳金和躲避申请营业执照相联系）⁹。还有未申报的小费和销

⁵ 一个实例是 Buehnat al. 所引用的全称题目为“影子经济和自制产品的活动”的论文。

⁶ 即“moonlightening”，意指兼有两份工作的人——校译者注。

⁷ 往往说不清楚它们是否被包含在“生产边界”（即 GDP 的定义），它们原则上由图 1 中的 1，2a 以及部分 2b 构成。

⁸ 注意：收受赃物（receiving）、造假、腐败与诸如此类属于再分配性的而不是生产性活动，因此，这类活动对 GDP 没有贡献。

⁹ 非法用工（主要是在建筑行业）是一种由雇主一方从事的黑色活动，就像“月光下的”或未申报的小费以及隐藏实物形式的收入一样，对于雇员和从事自由职业的个人业主（free-lance one man entrepreneurs）是很典型的（自我雇佣在 BE 中极为普遍，因为这非常容易对税务官员（taxman）予以隐瞒）

售额（如在农业领域）、以及未申报的实物工资¹⁰、雇员通过偷窃（小偷小摸）的“自我服务”方式“收受赃物”、填充开支账目、办公设备自用，等等。

图 1 中的 2b 部分，DIY——或者说“家庭自我服务部门”是“非官方的”，但并不被认为是“黑色的”¹¹。所谓的“邻里帮助”（neighbour help）不同于“月光下的活动”之处在于，在这种案例中，主要是为家庭成员或朋友工作，要么不付报酬，要么通过相互之间的安排来报答，用赠送礼物等方式来解决。同 2a 相反，大部分 2b 是有意地不被包括在 GDP 之中¹²。尤其是指没有报酬的非市场化的自用的（for own use）（如由“家庭妇女”所提供的）家庭生产¹³。2b 部分应该排除在 GDP 之外是因为将 2a 和 2b 这两类不同的事情混为一团没有意义¹⁴。因此，我们认为最好重点关注真正的黑色活动（即 2a），而不是还要考虑 2b。这也将有助于避免一直以来由“影子”和“黑色”相混淆造成的麻烦。

综上所述，“隐藏”、“影子”、“地下”或“黑色”这些术语之间存在着很多混淆，在第一和第二（影子）经济或黑色与白色经济之间也不容易明确界限，是由于在这些差别及标准之间没有一对一对应关系：

第一，生产边界（包含在 GDP 之中还是排除在 GDP 之外）；

第二，在从事生产的背景之下的“非法的”（illegal）或“非正规的”（irregular）的概念；

第三，自用的（非市场生产，常常是没有交易发生）或为了其他目的的（市场生产）；

第四，在（市场）交易（以货易货的或货币的交易，后一种情况是为了现金、债务或银行转账而从事的交易）的情况下所得到的那种报酬。

各种不同的“黑色”活动的唯一标准是共同的：它们缺乏可见性（lack of visibility），因为这些活动被刻意隐藏（隐藏黑色活动的动机在于逃避税收¹⁵以及逃避过多的管制）。这些活动肯定没有登记，从而仅能加以推算而了解（其精度程度会高于或低于实际情况）。

2. 有关原因和结果的几点说明

（1）影响影子经济规模的因素

就刚才提到的“转入地下”（immigration into underground）即转变为黑色经济（BE）的动机而言，很显然，黑色活动最重要的决定因素可能是同过度的管制相伴随的税负（还有税收的复杂性和对税收公平性的理解）（the complexity and perceived fairness of taxation）¹⁶。其他的因素还有失业、官方经济中工作时数的减少、官方与非官方经济之间的工资差距和（最

¹⁰ 正如在脚注 3 中所提到的那样，实物工资和未申报的小费以及销售均被明确地加以估算并被包含在官方 GDP 之中。

¹¹ 经济中的这一部分具有悠久的传统，可以追溯到市场还没有占主导地位而家庭并不仅是消费单位而且是生产单位的时代。因此，2b 在农村地区仍然是极为重要的，在农村，工作缺乏专业性和特定性仍然很盛行。进而言之，2b 有时可作为 2a 的替代物出现。

¹² 如果 2b 和家庭成员所有经过计算的产出的以及自己制作产品的活动都包含在 GDP 之中的话，那么实际上从经济学意义上来说，几乎所有成年人口都是积极人口，而且失业几乎完全被消除。基于同样的精神，来添加（从数量上看，肯定是含糊的而且仅仅是被推算出来的）微不足道的产值（如：准备饭菜，而且甚至是消遣的成果，如：园艺，采蘑菇等等），或者不付薪酬的私人户内的个人劳务（如：带小孩）对货币交易的价值，这同样是没有意义。这并不是说，这些活动对于福利没有贡献。这些考虑不仅对以上提到的按照 GDP 计算的表达 SE 产量的习惯提出了质疑（即使不重复计算的情况下，SE 也不能添加到“真正的”GDP 中去），而且对所谓的“无法观察到的收入假说”（unobserved incom hypothesis）提出了疑问（见下文）。

¹³ 例如：为住宅提供的清洁，装饰，维修的劳务，对小孩和病人的照料，为家庭成员的提供交通服务。个人拥有住宅是把自己使用住宅的无酬劳务从国民账户中的 GDP 排除出去的一般规则中的一个例外。

¹⁴ 我将在第一节中第 2 小节的第（2）点中重新讨论这一点。

¹⁵ 避税（tax evasion）并不是影子经济的必要标准（然而，一般来说，这样做很有可能是一种隐藏经济行为的动机，而被隐藏的经济行为是“黑色的”）。许多像 2b 一样的活动（如：“家庭妇女的”无报酬的劳务或者从偶然的工作中得到的可忽略不计的小额收入）无论怎么说是免税的，因此，避税很难成为一个问题。

¹⁶ 不仅仅是在产品市场，而且在劳动力市场。典型的规制是减少工作时间，或者提前退休年龄，最低工资，许可证（最低资格）要求。在某种程度上这些做法有助于对新来者进入劳动力市场设立障碍以减少竞争。

重要的)在(总的)劳动力成本和相当低的净工资之间(由于税收和社会保障缴纳金而造成)的“工资楔子”(wage wedge),这些因素为BE生产的顾客和供给者双方提供了进行互利的交易的机会。政府(例如:反“黑”活动)¹⁷的规模和效率无疑也是至关重要的。

(2) 研究影子经济的理由

说到研究影子经济的理由时,所谓的“无法观察到的收入假说”经常被人们提到,根据这一假说,(所谓的)忽视GDP中的影子经济可能导致官方对产出(GDP)¹⁸、通货膨胀、失业以及其他官方统计数字¹⁹的扭曲。而且,这反过来也会引起某种不稳定的反周期政策。另一个担忧当然是在于对税基的侵蚀(危及大规模再分配和福利制度的可行性),并且触发了一种持续高税收与为转向黑色经济提供了更多的激励二者之间的恶性循环的触发,通过这种方式使税基逐渐降低,这反过来再度要求增加税收,以便对转向BE的行为作出反应。

(3) 影子经济的效应

应该注意的是,总的来看,影子经济的效应并不总是不利的,因此通过一切可能的手段来反击它是不明智的。通过减少管制和降低税负将能更好地间接减少向影子经济转移的激励。而且人们公认,即使BE可以减少,它也不能(至少不能同等程度地)被官方经济所替代代替,而且影子经济会很好地促进伸缩性(promote flexibility),可以缓解某些社会紧张局势、还可以和平地发出抗议过高的税负和不透明的、不公正的、无效的而又过分复杂的税收制度和管制体制的信号²⁰(参见:尤其是Cassel(1984)对这种观点的论述以及在本文无法详细阐述的有关影子经济的理论)。

二、有关影子经济规模的统计估测方法

1、对某些间接方法的回顾和评论

这一部分,我们将简要回顾影子经济规模的统计估测法作为开始,并特别强调我们将以对货币方法和模型方法做一个简要概述和评论作为结束。(在本文中第二节第3小节和第4小节)我们再来更为详细地描述和论述货币法和模型法。这些方法通常被区分²¹如下:

(1) 直接法: 该法运用于:

a) 调查,如问卷调查有关人员他们是否已经或曾经参与了影子经济活动,假定他们回答属实,或者

b) 从税收审计中获得数据,例如:对已发现的未申报的税收所做的记录和BE操作造成的罚款,因此,数据仅代表着BE的一个相当宽泛的部分中的书面证据。²²

这些方法据说是仅仅为规模估测提供了最下限。

(2) 间接(或指数)法: 该法的导向是:

a) 非货币指标,例如:总量的国民账户、就业统计和物质投入(电力消耗)之间的差距或者“交易量除以GDP”²³比率的动态变化

¹⁷ 一个小的公共部门常常和一个小的SE相伴随(例如:日本,美国)。

¹⁸ 据说,例如:美国在20世纪70年代和80年代的生产率下降被夸大了。

¹⁹ 参见:第136页脚注⑨关于对这种不受欢迎预设的实质和辩解提出质疑的理由。

²⁰ 一个包含许多免税(漏洞)的复杂的税收体系——令人感到悲哀的是,德国曾经因此而取得了卓越的成就——会鼓励合法避税,逃税,甚至导致较低的纳税意识。

²¹ 这里,就这个术语而言,我们再一次不会有哪怕是微小程度(in no small measure)的混淆。

²² 很明显,调查以及所提及的官方记录将仅提供SE规模的最下限。

²³ Feige的交易法也可认为是货币法

b) 两种变异的货币指数

第一，天真的固定（现金-存款）比率法（a naïve fixed (cash-deposits) ratios approach）（参见：Cagan, Gutmann），以及

第二，建立在通货需求函数（currency demand function）估测基础上的 Tanzi 法

(3) 模型法

将影子经济作为一个变量在一个“结构方程模型”（a structural equation model, SEM）中进行估计，这是一种在统计学中同样著名的 LISREL 方法²⁴ 或影子经济模型法。在这种方法中，即使影子经济不直接被观测和度量，它（尽管不“显示”）却可视为一种由一组 q 个显变量（ η 是这些 x 变量和一个随机项的线性函数） x_1, \dots, x_q 导致的“潜”变量 η ， η 被视为是产生了其他显变量（或影响） y_1, \dots, y_p 。这些变量被称为指数。这就同样解释了这种模型为什么叫做“MIMIC”（Multiple-Indicators-Multiple-Causes）而不是叫作 LISREL。变量 η 的规范构建事实上代表了影子或者黑色经济的规模，这一事实是建立在模型中的 x 和 y 变量的属性的基础上的（例如 x_1 = 税负， x_2 = 管制的量，或 y_1 = 就业，等等）。

这里要对 2a 组的方法做一点简短的说明。2a 组的间接法定位某些统计数值间存在差异（如在宏观经济水平的收入和产出之间、或劳动力总量和官方记录的就业率²⁵之间），如果没有影子经济它们并不会相差太远，或者它们从“超出的”（通过某种度量标尺来判断）变量来推测影子经济的存在性，例如：货币流通或私人用户的电力消耗（或比产出多的“物质投入”）。²⁶ 这种方法存在的问题是，至少这些中的一个数值和其他比较（为了得到反映影子经济的“差异”或“不一致”）是很不可靠的，因为它们仅是从估算中得出。例如总体“可就业”人口与通过调查或薪金统计得到的官方报道的实际就业人数有差异。²⁷

比较收入法（雇员的补偿、财产收入等的总额）和国民账户中的支出法（私人消费，政府支出，投资和净出口）看起来更加令人困惑，因为这些 GDP 的初步估测不会公布²⁸，即便如此，此类差异度会比通常认为影子经济的规模小得多。此外这种方法错误和暗自假定在所有国民账户总量中没有测量误差，它同时假定差异意味着黑色活动。如果有什么差异，就是由于两种估计 GDP 的方法建立在不同的数据库基础上，这种差异基本来源于统计“误差或遗漏”。

另一种让人疑惑的方法是所谓的 Feige 交易方法，我们仅在此简单介绍。这种方法需要了解“交易量” $M_t V_t$ （货币为 M 和货币交易流通速度（而不是收入）），且将它和 GDP 比较。交易量和 GDP 的差额不仅是由于影子经济，也由于中间消费（原材料的购买）和纯金融交易。²⁹ 至于其他方法，还有“黄金时代”（没有影子经济），但很难建立实证。

影子经济，或者更大程度上说，BE 倾向于将其活动排除在政府之外，尤其是排除在税务当局之外，因而难以进行统计观测。因此，就定义而言，在统计上时无法观测的和没有被记录下来的行为。因此，测量影子经济是一种类似于“术语上的矛盾”（contradiction in terms）

²⁴ 线性独立结构关系（Linear Independent Structural Relationships）

²⁵ 这种方法在第 3 节中被称为“就业法”

²⁶ 这种物质投入法似乎仅是一种可运用于 SE 中家庭生产部分的方法。其他的投入指数，如 DIY（自我经营）商店的营业额经过讨论同样被认为很可能与 SE 有关（尤其是对于“月光下的活动，即兼职”而言）。

²⁷ 注意：有多于一个工作的工人将在薪金调查中被计算两次而在户调查中仅计算一次（就业统计）。

²⁸ 众所周知，三种“方法”，收入，支出和和生产法（利用诸如农业，工业，贸易等部门净产出的数字（或指数）可以（大体上）理想地和和在实质上归于同一结果（从而每一种方法都可以作为对其他方法做相互检验）。然而，在公开的 GDP 的数据中初始差异被取消了（“调整”或“协调”），因此，这是一种在统计上不可能采用的（所谓的）不一致规则。例如：由产出法和收入法计算的 GDP 差异在 1995 年的德国的是 0.8%（根据 Braakmann (2004)）。这些数字是通过提高生产法得出的数值 0.7% 和降低收入法得到的 0.1% 来平衡的。实际上，这种初始差异并没有公布，因而并不为统计值的使用者所知。

²⁹ 金融交易仅仅改变的是金融资产的结构（现存的财富），因此，它们对以 GDP 衡量的产出（即增加新的财富）没有贡献。

的某种东西。顺便说一下，Schneider & Enste (1999) 可能并不完全知道他们观点的荒谬性，这种观点注定是把他们这种“科学的”(!) 动力和动机描述成“一种探索未知的科学热情”，仿佛未知领域会是一个科学探寻的目标。

2、参照公理的评估法

在继续讨论其他的——或许更加可靠的——方法的优点和缺点之前，在这一点上，即思考一下作为一种合理且又被人们接受的方法应当达到的某些标准，似乎是有利的。接下来我们将列出初步收集的 6 条“公理”(axioms)，这正是某种方法的质量所要求的，如果对于好的统计操作是必不可少的话。它们是：

第一，有效性：这种方法应该充分包括目标概念——换言之，合理地反映所要测量的是什么——还要基本上建立在观测而非纯粹的推算基础上。如需假定，其数量应少些，至少应当可以做实证验证或起码要合理。

第二，数据基础：这一方法所要求的数据应当可靠、容易操作、足够详细且适于问卷（例如：对影子经济作为一个整体或者仅对地下经济），估算应当服从同一条思路。

第三，具体结果：这一方法应允许不同类型的影子经济活动之间可充分区分，其结果应有助于分析中采用（不仅提供影子经济规模的信息³⁰，而且在与其他变量的相关性，或者在影子经济的其他驱动力（或“原因”）相对重要性的基础上提供有关影子经济结构的信息。）

第四，理论基础：方法的理论基础应从微观或宏观经济模型中推导出来；例如：这几乎不能说是有关 MIMIC 方法的，在使用该方法时，只有变量的属性才会在模型中被加以考虑，从而我们可以相信我们正确地测度了我们所要测度的东西³¹。

第五，同其他方法一道进行交叉检验：最重要的是，该方法应该允许进行这种检验³²与合理性检验（为了进行估算，它还应给出一些有关拟合优度（goodness of fit）的数字指标（类似于统计学中的 R-平方、可决系数（the coefficient of determination in statistics）的某种东西）³³，最后，它应当遵从——

第六，良好的统计程序（职业操守）：这种方法应当很好地被记录下来，其结果一般来说至少是可以重复的。

可以肯定地说，任何当前正在使用的测量影子经济规模的方法远未达到这些标准。结果仅仅被说成是“从 x-方法中得到”或者更糟糕的是我们发现了类似于“基于 x 和 y 的方法”

³⁰ 在这个领域，一种通用的（而且在统计学上是不寻常的）做法旨在提供一个简单的数字，被假设为 SE 的“规模”，没有分解为 SE 的相关部分或者一个有关某些决定因素的相对重要性指标。存在着许多比 SE 规模更有趣的方面。在统计上，一般来说，孤立地选择一个数字只有有限的价值。

³¹ 我们或许还可以说，有关现金交易是更优类型或者甚至把这些操作的特征称之为“黑色”的简单假设对于支撑某种理论来说是太贫乏了，这是因为同样还存在着许多其他的可能激发“过多”货币需求的操作。

³² 对于此，3 种独立地估计 GDP 的方法是一个显著的例子。如上面所提到的，宏观经济恒等式所带来的一个后果在于不同的估计至少应该在原则上会得出相似的结果。这个平衡原理是国民账户框架中将估算加总的根本方法（这在其他情况下可能无法估计）。一般的想法是在不同的或者独立的方法之间进行调和，也许可以得到可信的数字。

³³ 对 R 平方信息进行估算（例如：Akaike 或 Schwarz 标准）使拟合优度和“吝啬的说明”（parsimonious specification）二者之间达到了平衡。更高的 R 平方值和更低的信息值是我们所需要的。

的说法。此外，这是一种罕见的现象，即更多的线索（hints）涉及估算而不只是数字（相对于 GDP 的规模的）数字以及所使用方法的名称。同其他方法得出的结果进行交叉检验，或明确地考虑被估计数字的可信度，这也是不多见的。最后，关于估测的准确性，人们了解得很少或者一无所知。所有这些无疑与统计学上的职业操守没有什么关系，而且由于影子经济自身的本质（尤其是 BE）是隐避的并且任何人都无法确定它的规模，因此，这种技术状况（this state of art）愈加令人感到悲叹。如果估计结果出来了，这些结果有可能都会被人相信，不管其所使用的方法是处在什么状态。Koch（2007）强调更多的是“从事估算的行当”（estimation business）应该充分意识到这一责任，这一点值得赞许。

3、货币法

运用货币变量作为度量影子经济规模的指数的方法——或更准确地说“黑色经济”——被称之为“货币”或“通货”方法。它们是更为广泛地被使用的方法。因此，在文献中，对它们的优点和缺点也比任何使用其他方法的案例讨论得更为彻底。普遍的看法是，“隐性”经济中的现金交易被人们更加偏爱，因为它们在簿记（bookkeeping）和银行转帐中无处查询³⁴。货币法在以下变量中存在：

第一，一个简单的变异（“固定比率法”）旨在把一些在货币总量（纸币，或称之为 M0, M1, M2 等等）或者在一些货币指标如通货-存款比（即 C/D 比率）所发生的显著变化（和基期 0 相比较）解释为黑色活动增加的指数。其隐含的假设是，如果影子经济（或 BE）没有上升，这种比率应该维持在基期水平上不变³⁵。经验证据显示出 C/D 比率波动很大。没有明显的理由可以解释为什么即使没有影子经济，C/D 比率在长期内会保持不变，而且更无法解释为什么一个 C/D 比率增加仅仅是由于影子经济所增加造成的。

第二，这一方法中更为复杂的变异，即“货币需求法”（CDM），由意大利人 Vito Tanzi 提出，它需要对通货需求函数进行统计估计，而在货币函数中，因变量（例如：通货 C 或者 C 的自然对数）是由利率，税负，管制强度的指数以及在一个回归方程框架中的其他解释变量来解释的。³⁶这一方法更有吸引力，这是因为我们不仅考察了数据，而且还能采用拟合这些变量的模型或多或少对这些变量做了适当的“解释”。这一方法允许把因变量中的部分变动（如：C, ln(C) 或 C/D-比率³⁷）视为或多或少由影子经济所造成，并且把对货币的额外（过度的或“非法”的）需求与正常预期的某种需求分离开来。这个方法依赖于假定，即给定基期解释变量值，回归函数应该给出正常的（或“合法”的对应于“非法”的）对货币（现金）的需求。

³⁴ 注意：对现金的偏好仅仅适用于需要隐藏的“黑色活动”。隐藏的 DIY 是没有意义的，它们在一些欧洲国家（如英国）是很流行的业余活动。这当然又提出了另一个问题：我们怎样才能运用一种货币（通货）法估计全部的 SE，而不仅仅是地下经济（SE 的 BE 部分）？

³⁵ 特别是在（Gutman 的）“货币存款比率”法中假定我们将有更多的 SE，以至于 C_t/D_t 比 C_0/D_0 更高。其中 C 代表通货，D 代表存款。

³⁶ Tanzi 的原始回归方程是 $\ln(C/M_2) = a_0 + a_1 \ln(1+T) + a_2 \ln(W/NI) + a_3 \ln(R) + a_4 \ln(Y) + e$ 其中 T 代表（加权的）平均税率，C 为货币，M2 为货币供给（=C+D），W/NI 表示工资和薪水在国民收入（财富的代理变量）中的比率，R 是对存款支付的利息（我们预期 $a_3 < 0$ ），Y 为真实人均国民收入以及 e 和往常一样，代表随机误差项（Tanzi, 1983）。在其他的 CDM 操作中，例如：对于澳大利亚来说，回归元如同税收系统的复杂性或政府管制的密集程度而被引入进来。参见：Schneider & Enste（1999, p. 30f.）。在其他的估计中（澳大利亚 1956-1991 年和 1956-1985 年）Schneider 对真实通货流通量的对数用被解释变量的滞后值、私人人均真实消费支出、利率、税负（直接和间接税）以及税收系统复杂性指数和管制程度指数（他在一个仅为 30 个年度观察变量的样本中利用至少 8 个回归元）。

³⁷ C 也可能是用人均绝对值表示的“真实”（即经过物价缩减指数调节的）通货。除 D 以外的分母有可能是 GDP、私人消费或类似的宏观经济变量。

“非法货币”应乘以 V ，即货币的收入流通速度（the income velocity of money）³⁸以得到所估计的影子经济的流转速度，进而反映影子经济的规模（假定影子经济中的 V 等于官方经济中的 V ）。

需要注意的是，比起简单的固定比率法，对预期的通货需求和额外的通货需求之间作出区分体现了某种显著的进步。还要注意的，这种回归法至少在原则上可以适当地应对来自标准的“货币法”的批评。比如，增长的黑色交易并不是通货需求背后的唯一驱动力。其他的决定因素正在使支付习惯、窖藏或某种“国际储备货币”的角色发生改变。应该通过在回归方程中增加适当的回归元，以便使检验这种论点的相关性成为可能。³⁹在这种情况下，仅有回归方程中未解释的残差（the unexplained residual variation）都会是 BE 活动增长的征兆。

第三，根据第三种方法，至少在大体上能做出较好的解释方法，黑色活动被假定同仅仅偏好于使用某种（更大）面额的纸币相联系（例如：100 美元或 100 欧元，200 欧元）。然而，大面额纸币数量不断增长也许可以归咎于通货膨胀，⁴⁰而且人们同样认为（Graf, 2009, pp.13-14），典型的黑色交易会建立建立在“正规的现金”（normal cash）例如“对于日常交易更常见的银行券”的基础之上。⁴¹这些争议再一次表明，测度影子经济的方法（或反倒有可能只是 SE 中的 BE 部分）建立在含糊不清的假定和众多合理论证的基础上。

第四，第三种方法有时根本就不被提及。相反，Schneider & Enste（1999）将以上所述的 Feige 交易方法命名为通货法的第三种变异。

作为一种规则，所有这些“货币法”依赖于这一假定，即“黑色”活动大部分是用现金来交易的，从而可以隐瞒税收机构且不留任何可以观察到的痕迹。作为影子经济（或者说 BE）的指数，现金货币（cash money）总额似乎是一个好的选择。幸运的是，它同样是一个已经可以获得的统计数字。

同所有货币法（同样还有其他的方法⁴²）相联系的一个问题是，它们需要一个没有影子经济的“黄金时代”（它可以用作“基”期 0（“base”-period 0））。⁴³然而，当我们承认无法确定影子经济应当怎样被测度出来时，那么，在实践中，似乎难以确认这样一个没有影子经济的阶段。

CDM 法的理由在于，相对于回归元取“没有影子经济情形”的水平时，过多的通货需求会仅仅因“黑色”经济活动而产生。如果实际通货需求（作为因变量）低于回归元在基期水平上的期望水平时，这个逻辑不可避免地将陷入困境，这是因为难以想象一个过多的负的通货需求的结果的是一个负的黑色经济规模。正如图 2 所示，这并非完全不可能。2001 年底德国的大量窖藏现金在临近引入欧元时却被蒸发掉了便是一例。

除此之外，CDM 法隐含地提出了这样一个未被接受的假设，即影子经济是完全地和唯一地由回归因子决定，以及这些回归因子可以正确地捕获住，而且它们可以把所有的成因反映出来。由于种种原因，这个假设是站不住脚的。

第一，这些作为回归中的因变量来影响货币变量的因素（如通货）总的来说只能相当含糊地被确定，就像是在上面提到的“税收系统的复杂性”或“政府管制的程度”这类因素一

³⁸ V 被定义为在一段时间内（例如一年）“流量”指标（收入）并被定义为某一时点上的存量指标（货币）的一个比率。因此，当一个存量“不合法的货币”需要转换成一个流量（用 GDP 表示的 SE 的规模），拟必须将这个存量乘以 V 。

³⁹ 与上面提到的管制总额、税负等变量不同，这些变量并不被当成 BE 的决定因素或相关因素，因此，这些回归元的取值（如：外国的通货需求等），不应当在基期值基础上设置，而应当基于它们的当期值以得到“过多的”（由此而引致了 BE）通货需求。

⁴⁰ 在通货膨胀的过程中，大面额纸币不再和它们当初那样大。并且纸币 100\$ 美元或 200 德国马克（或者在德国马克的案例中曾经发生的那样）经常在其他国家被用于财富储藏。

⁴¹ Graf 的论点主要是：“大面额的货币可能更容易被察觉。”

⁴² 包括交易方法

⁴³ 这样一种“无 SE 情形”是需要的，进而决定“正规的”现金货币需求，该需求被假定为在基期 0 水平上当税负和其他影响通常维持不变时占主导地位。

样。它们应当被转换成在操作上明确而又可观测的变量（例如：相关指令的数目，立法行为等），这些变量可用作回归元 x_1, \dots, x_q 。很难证明回归元能够实际上足以代表这些因素。

第二，（已观测到的）因变量 y 很显然不是（未观测到的）目标概念，如 z ，即影子经济的规模。因此，这个方法依赖于一种设法将 y 和 z 联系在一起的理论，而认识到这一点是很重要的，即这一理论并没被回归（尽管良好的 y 表示为 x_1, \dots, x_q 的函数可以拟合数据）所证实。一个很好拟合的通货需求方程绝不是支持某种主张通货需求是影子经济的一个指数的理论的证据。

第三，除此之外，有许多 CDM 实证运用的例子。在这些例子中，不同的通货需求回归方程所做的设定导致了一个大体上相等的“拟合优度”（用经过调整的 R 平方来度量）。进而言之，回归元的参数可能不稳定，而且明显与影子经济无关的回归方程可能产生一个相等的拟合优度（an equally good fit）。

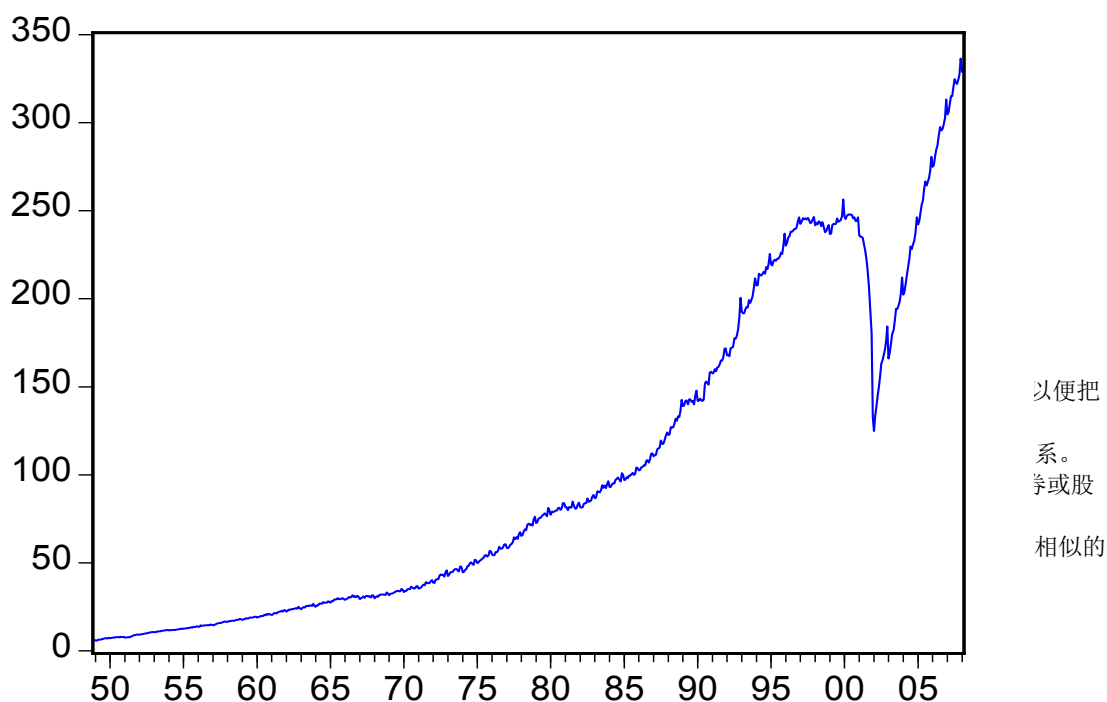
如果通货需求函数因为它处于 CDM 的中心而有什么不同的话，那就是它不是唯一的决定因素，而且从通货统计值中推算影子经济的规模（仅仅用作 SE 的指数）也并不是那么直截了当。

值得注意的是，以上 3 个论点中的每一个论点都可以同样运用于（甚至更有可能）MIMIC 方法，另一种度量影子经济规模的统计方法将在第 4 小节中予以讨论。

在实践中，其他了采用 CDM 的方法同样需要对 SE 中的货币流通速度 V 进行估计。⁴⁴这个流通速度似乎再一次成了不容易度量的某种东西（假如所有的意图均在于测量影子经济而不是假定影子经济的话）。

然而，反对 CDM 或者货币法的其他变种的反对意见在于，现金交易的增加对于表明黑色活动的增加来说，既不是必要条件，也不是充分条件，而通货需求受到更多因素的影响而不仅仅是受到影子经济规模增加的影响。其中最为确定的一个因素是这一事实，即一些货币如美元或与此同时日益强大的欧元是被当作“国际储备货币”⁴⁵或者作为一种比以本国货币来储存财富更好的防止财富在通货膨胀中贬值的媒介。⁴⁶在以前，这同样适用于德国马克（DM），它在一些东欧国家被作为“平行货币”（parallel currency）而被持有。这种国际窖藏行为在 2001 年最为凸显，当时达到了令人难以置信的程度，即以窖藏形式持有的德国马克在德国和国外都明显地消失了（参见：图 2 1949-2008 年间流通中的德国马克纸币）⁴⁷以便为即将到来的 2002 年 1 月欧元替代马克做好准备。

图 2：流通中的德国通货（仅纸币）（直到 2001 年为马克，其后为欧元）



以便把
系。
等或股
相似的

注：单位为 10 亿德国马克。

在这一过程结束的 2001 年的最后几个月，德国马克纸币量消失到仅有两年前的 1999 年水平的一半，从那以后才逐渐在 2002 年的最初几个月达到了其正常水平。略看一下货币总量 M1（包含活期存款加上通货）显示，德国马克纸币显然被转换成了存款（Graf, 2009）。有趣并令人惊讶的是，这样一场在德国货币形势中意义深远且几乎是革命性的变化，在那些表面上以 CDM 为基础对影子经济进行估算的学者们所提供的有关德国部门规模的估测中，竟然没有留下任何印迹。⁴⁸

对这种估计进行可信性检验相当罕见。⁴⁹但是，还是存在着建立在可信性考虑基础上的理由让人们相信，这些估测被夸大其辞了。⁵⁰这将在第三节中加以说明。

4、基于模型的方法：黑色经济作为黑箱的产出

在这一节中，我们再次与被认为可以度量影子经济的方法相遇，这种方法依赖于对统计数据中的因果关系作出解释的有效性，然而，在数学上要采用更为精确的方法。在这里所提到的方法使用了不同的名称。⁵¹在接下来的部分，它作为 MIMIC 而被提及。从纯粹统计学的观点来看，这种方法肯定是相当有趣的。这个方法是否在估计影子经济规模的特定领域中前景广阔，仍有待观察。正如已经指出的那样，大多数针对 CDM 的反对意见（与可观测的和不可观测的变量相联系的基础理论，这些变量是否能真正反映了人们想要测量的东西，对有可能使拟合数据达到同等优度等诸如此类的方程的种类尽可能作出详细说明）也适用于 MIMIC。此外，MIMIC 模型方法经常和 CDM 结合起来使用。其理由是为了估计 SE 的时间序列的（绝对）水平，而不是 SE 潜在变量的水平变动，通过采用从另一种方法中选取的对 SE 做点估计（即对时间上的一个点进行估算）的方法使模型得以“校准”。在这里，由于在习惯上依赖于 CDM 的缘故，并不令人感到奇怪的是，从 MIMIC 得到的实证结果和使用 CDM 得到的结果十分类似。⁵²很明显，这样一来，结果上的相似性并不能证明其中之一或两种方法的有效性。没有哪一种方法能够用作对其他方法做交叉检验（cross-check）。

MIMIC 的基本思想很容易掌握，这一基本思想在图 3 中被展现了出来。

⁴⁸ 没有人会将德国马克纸币的减少解释成一个“黑色活动显著减少的信号”（Graf, 2009, p. 11）。Graf 的观点如下：“如果这一规模的现金量能够在私人消费支出没有反应的前提下蒸发的话，先前的现金增长同样无法被视为德国经济部门黑色活动有更大规模交易量的证据。”（p. 14）。

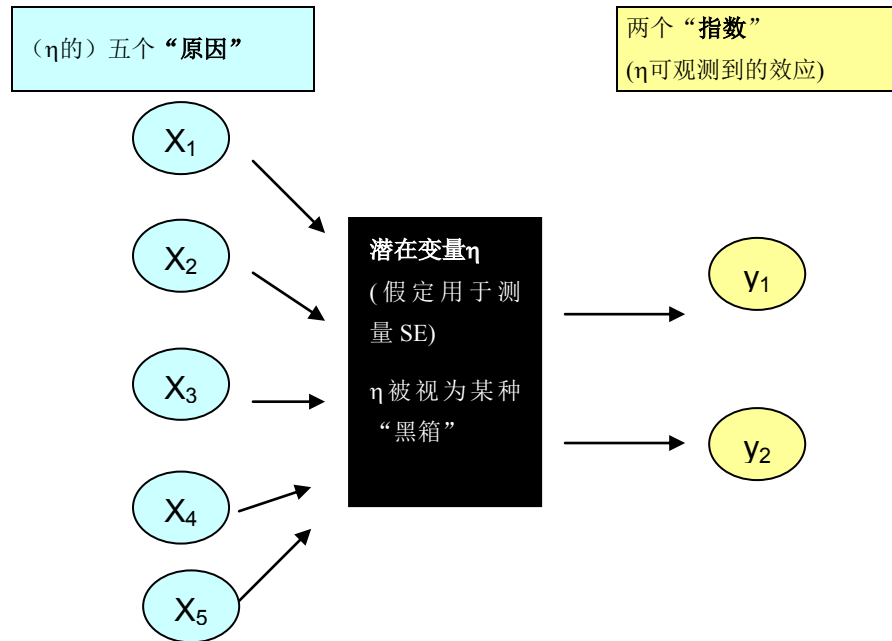
⁴⁹ 正如上面所提到的，在用其他方法度量 SE 的案例中也不常见。

⁵⁰ 它们由 Graf（2009）或 Koch（2007）提出。例如，Graf（2009）认为，尽管货币流通日益增加，在德国马克制度时代的大部分时期，黑色活动起着极小的甚至减弱的作用。

⁵¹ 除 MIMIC 以外，还有 DYMIMIC（Dynamic MIMIC）或者 DSEM（Dynamic Structural Equation Model）。在统计上更喜欢把上面提到的术语叫作 LISREL。

⁵² 这是 MIMIC 和 CDM 在大多数实证运用中相对相近的理由，这一点在 Buehn et al. (2009) 中明确得到认可。

图 3：在一个 MIMIC 模型中的影子经济概念化的例子¹⁾ (Martin Eiglsperger 的一个模型)*



原因²⁾: x_1 = VAT/GDP 比率, x_2 = 工作时间, x_3 = 可支配收入, x_4 = 边际税率 (所得税), x_5 = 税收与社会缴纳税金 (social contributions) 总额	指数³⁾: y_1 = 增加值 (净产出, 对 GDP 的贡献) (部门 S.11 和 S.12 (用来使模型标准化)), y_2 = 就业参与率
其他候选“原因”: (实际和可以被察觉的) 税负, 管制的强度, 相对于 SE 而言官方经济中的工资和价格, 失业率, 通货膨胀率, 某些哑变量 ³⁾	其他候选“指数”: 国内流通的货币, 官方经济产出 (的增长率), 正规部门中 (减少的) 就业, 税收收入的 (减少), 等等

注: 1) 多重指标——多重原因; 2) Eiglsperger 模型中的案例; 3) 例举诸如德国统一或从德国马克向欧元转换等事件; 4)* 同 Eiglsperger (2004) 相比较, 还可以参见 Buehn et al. (2009) 分析 MIMIC 模型的另一个例子。

影子经济虽然没有被准确地定义, 却可当成一个 (不能观测到的) 潜变量 η , 它被定义成它和某些显 (可观测的, 统计上可度量的) 变量 x 和 y 的线性或随机关系。这些 x 变量被认为是 η 的原因 (在这个意义上的因变量, 或者在一个回归模型中“解释”因变量 η 的“回归元”), 由于高的 η 值和高的 x 值相伴随。⁵³ 与此同时, 变量 η 代表影子经济反过来带来某些效应, 使得 y 会发生 (或者“导致”⁵⁴ 它们会用因素分析来理解“因素”)。 y 变量因而可被视为可见的影子经济的显示或“指标” (例如: 现金持有量, 生产的增长率)。

因此, 该模型结合了部分类似回归, 它们将 q 个不同的 x 变量和单一的 η 变量联系成如下形式 (T 个观测中的第 t 个方程, $t=1, \dots, T$)

$$\eta = \gamma'x + \zeta \quad (1)$$

另一方面, 一部分类似的因素分析⁵⁵

⁵³ 这一思想似乎是支配着对“原因”变量 x 以及诸如此类的变量进行选择的规则, 例如: 官方经济中的 (实际的和接受的) 税收水平、就业, 工资和价格等等。

⁵⁴ 注意: 这个 y 变量在统计上和 η 关联的方式正如 η 与 y 变量关联的方式是一样的。

⁵⁵ 在因素分析中, 我们通常有一个涉及诸多潜变量 (“因素”) $\eta_1, \eta_2, \dots, \eta_m$ 的数字。同样一个 MIMIC 模型也可能拥有多于一个的 η 变量, 其中不像因素分析中的 η_1 和 η_2 会和其他的 (可能仅部分重合) x 和 y 变量

$$\mathbf{y} = \eta\boldsymbol{\lambda} + \boldsymbol{\varepsilon} \quad (2)$$

在方程(1)中 $q \times 1$ 的列向量 \mathbf{x} 包含了 t 期观测值的 q 个“原因”，即 x_{1t}, \dots, x_{qt} 以及 $1 \times q$ 的行向量 $\boldsymbol{\gamma}'$ 包含了相应的回归参数。在(2)中 \mathbf{y} 是一列有 p 个指数的向量 y_{1t}, \dots, y_{pt} 。除此之外，在向量 $\boldsymbol{\zeta}$ 和 $\boldsymbol{\varepsilon}$ 中分别有两组不同的 (q 和 p) 随机变量，关于随机变量 $\boldsymbol{\zeta}$ 和 $\boldsymbol{\varepsilon}$ 的方差以及它们之间的协方差的假设是估计参数向量 $\boldsymbol{\gamma}$ 和 $\boldsymbol{\lambda}$ 的关键。⁵⁶

因此，MIMIC 根据影子经济似乎可以被估测的性质而合并了两个方程组（因素分析和回归分析的混合）。然而，由于系统复杂性以及很快将要做出解释的诸原因，影子经济作为某种黑箱的产出而出现。在以下事实为给定条件下，影子经济似乎更像是纯粹的人造物 (mere artefact)：

第一，影子经济（即变量 η ）首先仅是形式上（数学的）“构造”。不存在普遍接受的有关影子经济的定义，也不存在建立以这两个方程组作为根基的经济理论基础上的关于影子经济的模型。如果有这样的定义，它应该有可能通过对它的适当量化的部分 (aggregating its properly quantified components) 进行加总来估算影子经济的规模。然而，这并不是影子经济被度量的方式。MIMIC 仅提供了一个不能表明它怎样被分解成各个部分的总体画面。

第二，使我们能够相信的是，这个在统计上得到的潜变量 η 确实是影子经济，而且我们已经有效地度量了影子经济的规模，这样做仅仅是处于这个想法，即选取某种变量，以便在模型中用作“原因”和“指数”（结果）的变量，这些变量似乎与通常由影子经济所表示的含义有着某种相关联的东西。

第三，不存在针对选择合适变量和针对模型以及是在 x 变量组中还是在 y 变量组中设定一个变量的简单法则。尤其是，我们在两边，即“原因”和“指数”（例如：反映收入发展的变化）。发现了十分相似并且在结果上非常相关的变量。⁵⁷当然，作为规则，在对统计模型中所采用的变量做出决策时，预先决定它将如何拟合数据。这一规则在更大的规模上运用到甚至像 MIMIC 一样一个的模型中似乎并不是不可能的。⁵⁸

第四，如前面提到的，MIMIC 没有提供影子经济绝对规模的数值，而仅仅是关于其原因或效应和系数相对重要性（相关性）的信息，这些信息允许对影子经济的动态进行估算。为了给出用绝对项 (absolute terms) 表示的时间序列，这些估算必须从运用另一种方法（当然，更偏好于运用相同的数据）中得到的某个基准（例如：CDM）相联系。

第五，除此之外，这个模型要求研究者做许多的假定和决定，这些假定和决定会要求某个黑箱的这种结果的可重复性 (reproducibility of results) 提出质疑。这里把一些有关的步骤罗列如下：

(1) 模型要求平稳的（不存在趋势的） x 和 y 变量序列（或者 x 和 y 变量的合理转换），因此，数据必须通过差分或通过运用诸如 RegARIMA 模型的时间序列回归模型⁵⁹来“去掉趋势” (de-trended)。所以，在一个复杂的对数据进行“提前检验” (pre-testing) 的过程中，这个被用作 x 和 y 变量的时间序列必须做非平稳检验（以检验单位根）和协整检验，而且并不是 x 或 y 的绝对值，而是它们的残差，被称为“创新” (innovations)，并且被作为真正的“原因”和“指数”来对待。⁶⁰

相关。方程 2 作为 MIMIC 模型里的度量模型同样为人们所熟知。

⁵⁶ 参考 Buehn et al. (2009) 并且在这一部分完全相同 参见 Buehn & Eichler (2009) 中更详细的内容。

⁵⁷ 例如：家庭的税负被放在 x 这边，而预算部门的税收收入放在 y 这边，或足以让人感到奇怪的是，失业率 U 可能出现在“原因”这边，而它的对应部分就业参与率 $E=1-U$ 被放在“指数”这边。

⁵⁸ X 变量因回归模型的性质而与 η 相关联。另一方面，“因素” η 必然地和 y 变量相关联。因此，如果每一个 x 和每一个 y 有充分的关联性，应该可能找到合适的中间变量 η 来很好地“解释”这些相关性。

⁵⁹ 参见：Eiglsperger (2003) 关于这一问题的详细讨论。

⁶⁰ 将 SE 的增长或衰落看作是由或多或少也许在概念上同 SE 有关的变量中的某种（随机的）“创新”所产生的，这似乎让人觉得很奇怪。然而，这在本质上看起来是当 MIMIC 方法被采用时会有什么事情发生。更需要进一步注意的是，模型中的变量可能以不同的方式处理，例如：人们可能仅对某些而不是全部变量的

(2) 正如其他的统计方法一样, MIMIC 模型要求对异常值(极值)进行有效的处理, 亦即对所确认的异常值进行识别和删除。

(3) 由于技术原因(因为对参数向量 γ 和 λ 的估计等于是一道“特征值”难题(amounts to an eigenvalue problem)), λ 系数的其中之一应该被“标准化”以便引入一个先验值(a priori value), 例如: 通过设定 $\lambda_1 = 1$ 。⁶¹

因此, MIMIC 要求采取相当一部分方法上的决策, 当不同的研究者会通过使用 MIMIC 获益时, 就将会提出结果的可重复性的问题。

总之, MIMIC 并不是毫无保留地值得赞许。虽然作为一种统计方法来说它非常复杂, 但似乎以这种方法度量影子经济并不能证明它比其他方法明显地更为合理。似乎并没有很多有关影子经济的经济理论突发式地采用 MIMIC 方法对影子经济规模进行估计。然而, 给定在对“影子”或“黑色”经济活动下定义面临着众所周知的困难, 同时给定这样一个事实即这类活动谨慎地被取消(隐藏), 以至于 SE 和 BE 只能(在最佳的情况下)“间接地”被度量, 就这一点而论, 这个有关“潜”变量或纯粹统计学意义上人造物(purely statistical artefact)的观点并不是完全不合适的。⁶²从这个观点来看, MIMI 似乎并不比任何其他方法更招致反对意见。

最后, 当所度量的结果与以上第二节第 2 小节所阐述的(公认为所需要的)在统计上进行理想度量的公理相悖时, 我们离显著而有效地度量影子经济的方法仍然还有相当长的距离。没有任何一种方法能够满足所有这些公理, 或者仅仅只能满足其中的大部分定理。这一状况很显然无法令人满意。当判定这些实证估计的价值时, 这种状况让我们多少感到无望。⁶³接下来在本文的最后一部分, 我们也许至少要确认不可接受的估计值。

三、对已发表的有关影子经济规模的计量统计的评论

若要查找一些关于黑色经济活动的数值信息, 人们不可避免地会同 Friedrich Schneider 教授、博士不期而遇。学术估计(或“猜-算”(guess-estimation))的领域——至少在德国——被已经奉献了几十年的 Schneider 所主导。如果精确地说, 他的全部学术生涯并不仅仅局限在这个题目上的话。⁶⁴他几乎垄断了这一领域, 并且因为他为全世界其他国家(至少 70 个)提供了巨大数额的估算而闻名, 以至于随着时间的推移, 他跃升为有点像这类行当中的教皇(a kind of Pope in these affairs)。尽管他写了大量的学术文章, 他却并没有就他所使用的方法在诸多细节的沟通方面动过脑筋。正因为这些理由, 他只是到最近才成为被批评的目标。不仅由于他在方法上不是基于充分的细节, 并且还保留了某些模糊之处, 而且他的结果看起来

情况下进行一阶差分, 而且“差分”的序列和 RegARIMA 残差混在一起使用。

⁶¹ 然而, 据认为“选择潜变量以便固定潜变量的规模, 这并不影响结果”, Buehn & Eichler (2009), 因此以上所讨论的问题看起来并不重要。

⁶² 使用因素分析法和潜变量(即因素分析背景下的“因素”)概念的一个成功的例子是对“智力”的测量。确切地定义“智力”并不容易, 并且通询问人们他们认为自己是否聪明(还可能问崇高到何种程度)来测量崇高, 这毫无意义。因此, 原则上通过将其作为一个潜变量(统计意义上的)来建构模型, 以便使一个模糊的概念量化这样的做法不受到反对(be frowned upon)。有时, 这或许是唯一有意义并可以获得的方法。一切都取决于模型的细节。

⁶³ 这正是公理的功能, 即给出一种方法优于另一种方法的理由, 由此对不同方法得到的结果进行评估。

⁶⁴ Schneider 看来像是一位杰出专家, 他从一篇论文中产生出至少有一打论文, 并由此达到了一个惊人的出版记录。我从没有见过还有人拥有对发表的大量论文上能够精雕细琢达到这样的程度的习惯, 这些论文相互之间仅仅在极其细微的方面存在着差异。在我看来, 这是这些年来在大学中盛行的资格制度(专业职称晋升标准, career criteria)真正的不良后果之一。从这些环境中所产生的另一种异常现象——对此, 我也曾在武汉研讨会期间评论过——是做一场希望因高深的数学而无人能理解的讲座, 借此希望给观众留下印象而已。在经济学和统计学中, 用数学把戏(wizardry)来炫耀, 并且为在那些对于其他人来说兴趣有限或者也许实际上根本就不相干的事物中拥有自己的专长而引以为自豪, 现在已经变得越来越流行。

显著地有点上偏 (biased upwards)。

在表 1 中我们列出了他提供的对一些欧洲国家和亚洲国家的估计。有趣的是，人们看到不同的方法如何带来了千差万别的结果。看起来可信的是，南欧国家诸如希腊或意大利⁶⁵可能有着相当大规模的影子经济（约占 GDP 的 25%至 30%），而且这一点（在更小的程度上）对于斯堪的纳维亚国家同样是成立的，这是因为这些国家的高税率所致，而德国仅仅只有一个适度规模（moderate size）的影子经济部门。

表 1：对 GDP 百分比中影子经济规模的某些估计（以 F. Schneider 的研究为依据）

欧洲		亚洲	
国家	规模*	国家（地区）	规模（年）
希腊	29.6, 28	印度	22.4 (1989/90)
意大利	26.0, 24	中国台湾	16.5 (1990/93)
西班牙	22.4, 21	中国香港	13 (1989/90)
瑞典	18.6, 18	泰国	71 (1989/90)
德国	13.5, 15	南朝	38 (1989/90)
瑞士	6.7, 9	菲律宾	50 (1989/90)

注：* 第一个数字为 1994-1995 年间的的数据，第二个数字为 2005 年的数据。

在文献中，我们看到仅有非常少的人把从不同方法得到的估计数据所进行的比较运用于对同一国家和同一时期。如果确实有这样的研究，其结果似乎相差很大，例如：（按照 Schneider 的说法），在 1989-1990 年的墨西哥：电力投入法得到 49%，而货币法（CDM）得到 33%。在另一项研究中，Schneider 对一组 5 个国家⁶⁶1970-1990 年间度量影子经济规模的平均值进行了比较，它们的值分别在 3.1%（普查方法）和 7.9%（MIMIC）到不低于 24.4%（通过对官方就业数字进行比较以及对劳动力总量进行推算而得到）。当不同的研究者用同一种方法来估计时，我们同样会看到极大的差异。⁶⁷

正如上面已经提到的那样，一些对影子经济的估计（大部分是 Schneider 基于 CDM 的估计）常常不被人们怀疑为夸大其辞了。基本上有两种方法可从使得这种怀疑得以体现，这种怀疑可以在有关这方面的文献中找到，即：

（1）如果对已观察到的货币总量发展不可能做出其他解释的话，则通过比较货币（尤其是现金）、消费或其他变量的时间序列以及它们各自的增长率，至少被认为同样是可信的。

（2）如果对数字所做的估算是认真的，而且它们照原样被接受，以至于估算看起来不合格，那么，用数值来显示含义和结果既极端又荒谬。

第一种方法所固有的困境，即仅仅从纯粹统计学的描述性方法做出推断（而不是对一个计量经济学模型进行估算）是有困难的，例如：如果得出有关各种不同影响作用于现金需求

⁶⁵ 尽管 Schneider 报告其估计为大约 26%，意大利官方对意大利的估计仅仅约 17%（参见：Koch, 2008, p. 93）。Koch 还引述了欧盟委员会所表达的关于 Schneider 对欧洲国家的估计（p. 95）所持的保留意见（即对上偏表示怀疑）。

⁶⁶ 加拿大，美国，德国，英国和意大利。

⁶⁷ 明显的例证是，据 Schneider 的研究，在苏联解体后，SE 的增长速度为：俄罗斯从 14.7%（1989/1990 年）上升到 41%（1994/1995 年），格鲁及亚甚至从 24.9%增长到 63%。

（像影子经济那样，改变支付习惯，窖藏或国外货币需求等等）的确定而又清晰的结论并不是不可能的话。这尤其适用于当这类要素有可能在不同方向上起作用，而不是一种要素被另一种要素所抵消。当缺乏一种允许仅仅从观察中就能明确地推断出一个因果顺序的方法时，⁶⁸，总是会有可能对同样的统计数值给出相互矛盾的解释。我们完全可以认为上升的（或下降的）“现金比率”就像是同私人消费支出相关的通货或 DIY 商店的销售额（the turnover of DiY-shops）有关的现金一样，都应当等同于更多（或更少的）影子经济。然而，同样有可能认为，如果现金比率随时间而下降，影子经济可能照样增长⁶⁹，因为存在着其他处在适当位置上的因素在发挥其功能以抵消影子经济的影响，以至于如果这些因素没有产生效力，现金比率的下降幅度会更小。因此，取决于假设某种以统计学中（数值的）观察为基础的因果机制的方法论难题就在于，任何机制都难以证实它是唯一的一个产生了观察结果的机制。

批评某种经验估计的第二种方法更有希望。它涉及表明估计值的含义是如何难以容忍，假如它会是正确的话。例如：Koch（2007）试图表明，如果 Schneider 关于 2004 至 2007 年德国影子经济的总量值为 3500 亿欧元是正确的话，这就会意味着：

（1）只要失业人员曾经参与这类隐性活动，他们的人均收入肯定会大幅度高于平均水平，而且花费在这种工作上的时间应该是常规工作时间的 26%，这看起来对每一个人都极不可能，同样地

（2）如果每个积极的（在就业统计意义而言）劳动人口以每小时 10 欧元的平均工资参与影子经济，这将会是总额达到 55.5% 的额外工作小时数（即超过正常的平均工作时间的追加部分）；

（3）如果对建筑部门影子经济生产的估计是正确的话，那么，这个部门的“黑色增加值”（black value added）肯定是“白色增加值”的 160%，这同样是难以想像的。

还有更多的这一类对估计值大体上进行检验的例子，这些估计值似乎在很大程度上显得有些上偏了。

应当充分表明的是，尽管目前对所有使用中的估计方法都有着一定程度的怀疑，但幸运的是，这足以说明至少可能识别了并排除了不合适的方法。从根本上说，在这种形势下，什么才是我们应当做的？首先，就有关定义问题以及停止对“影子”与“黑色”二者之间的混淆达成共识；其次，开始公布有关方法的细节、在估算过程中的起支撑作用的数据（underlying data）和“中间产品”，以便做出透明的假设和有限的结论。这一点同样可运用于官方统计。很难理解的是，某一个估计值好到足以进入官方 GDP 值，但与此同时仍然被怀疑到不能单独公布。我们还应当放弃那些以为建立在以不充分的文献为依据的方法基础上的数字没有问题的想法。最后，当被度量的东西是什么以及它是怎样得到的这类问题是值得怀疑的话，那么，也难以从拥有大量的估计值当中看出其中的益处。

（罗连发 译，马颖 校）

⁶⁸ 从哲学的意义上说，存在着充分的理由怀疑到底会不会一定有这样一种方法。如果关于因果关系的假定对于度量某种事物的统计方法是关键性的话，这一方法总是被认为是脆弱的，但这种看法是正确的，因此作为一条规则——而且还像我们在第二节第 2 小节中所陈述的第一条定理那样——某种基于观测值的方法应该优越于一种建立在（因果）假设基础上的方法。

⁶⁹ 参见：Graf（2009, p. 10），以便参考其他作者论点中的这个逻辑。

缩写

1. ARIMA auto regression integrated moving average model (RegARIMA when combined with a regression)
2. CDM = currency demand function method
3. DIY = do it yourself
4. DM = Deutsche Mark, former currency in Germany

参考文献

1. Buehn, Andreas & Eichler, 2009, "*Shadow Economy and Do-it-yourself Activities: The German Case*", in: Journal of Institutional and Theoretical Economics
2. Cassel, Dieter, 1984, "*The Growing Shadow Economy: Implications for Stabilisation Policy*", in: Intereconomics, Sept./Oct. 1984, p. 219 - 225
3. Eiglsperger, Martin, 2003, "*Quantitative Analyse der Untergrundwirtschaft mit einem linearen Strukturmodell*", Möglichkeiten und Grenzen am Beispiel der Bundesrepublik Deutschland 1976 bis 1989, Diss. Bamberg 2003, Frankfurt/Main (Peter Lang) 2004
4. Graf, Gerhard, 2009, "*Some Stylised Fact about the Connection between Cash and Black Economies in Germany*", unpublished paper
5. Koch, Walter, 2007, "*Zum Umfang der Schwarzarbeit in Deutschland*", in: List Forum, vol. 33, p. 153 - 172.
6. Koch, Walter, 2008, "*Sisyphusarbeiten - Untersuchungen zu Schattenwirtschaft und Schwarzarbeit*", in: List Forum, vol. 34, p. 81 - 101.
7. Schneider, Friedrich & Enste Dominik, 1999, "*Shadow Economies Around the World - Size, Causes, and Consequences*", Lectiones Jenenses (Working Papers of the Max Planck Institute for Research into Economic Systems)
8. Tanzi, Vito, 1983, "*The Underground Economy in the United States: Annual Estimates*", 1930 - 80, IMF Staff Papers, 30, p. 283 - 305.