



XÂY DỰNG KHUNG CHẤT LƯỢNG VÀ CÁC CHỈ TIÊU CHẤT LƯỢNG TẠI CỤC THỐNG KÊ LAO ĐỘNG

Michael Horrigan, Polly Phipps, Scott Flicker

Cục Thống kê Lao động, 2 Massachusetts Ave. N.E., Washington, DC 20212

Tóm tắt:

Trong hai năm qua, Văn phòng Giá cả và Điều kiện sống (OPLC) của Cục Thống kê Lao động Hoa Kỳ đã thực hiện một dự án để xác định khung chất lượng cho bốn chương trình giá. Sáng kiến này bắt đầu với việc xem xét lại các khái niệm và khung chất lượng; chúng tôi tóm tắt những phát hiện từ nỗ lực đó, phân biệt giữa chất lượng sản phẩm và chất lượng quy trình và xem xét các biện pháp chất lượng liên quan. Sau đó, các nhà quản lý chương trình OPLC đã báo cáo các biện pháp và biến số mà họ theo dõi như một phần nỗ lực của họ để đánh giá chất lượng. Điều nổi lên là sự phân chia giữa tầm quan trọng tương đối lớn hơn trong tài liệu chất lượng về chất lượng đầu ra và giảm tổng sai số điều tra và theo dõi cân bằng hơn cả về các biện pháp chất lượng đầu ra và các biện pháp vận hành chi tiết về chất lượng của quy trình tác nghiệp điều tra - đầu vào để phát triển đầu ra chất lượng cao. Dựa trên nhận thức này, OPLC đã phát triển một khung đồng thuận cho bốn chương trình kết hợp cả các biện pháp chất lượng đầu vào (quy trình) và đầu ra (sản phẩm). Những gì nổi lên từ nỗ lực này là một khung áp dụng các phương pháp kiểm soát quy trình thống kê hỗn hợp đến chất lượng với các khung quản lý chất lượng truyền thống được sử dụng trên nhiều cơ quan thống kê.

Từ khóa: Tổng sai số điều tra, quản lý chất lượng, kiểm soát quy trình thống kê.

1. Giới thiệu

Cục Thống kê Lao động Hoa Kỳ (BLS) là cơ quan tìm hiểu thực tế chính trong Chính phủ Liên bang đối với các lĩnh vực rộng lớn về kinh tế lao động và điều kiện lao động. Đầu ra quan trọng có thể được phân loại thành bốn khía cạnh - giá cả; việc làm và thất nghiệp; bồi thường và điều kiện làm việc; và năng suất - với nhiều chương trình trong mỗi khía cạnh. Đã có nhiều công trình đáng kể trong nhiều thập kỷ trở lại đây trong việc phát triển các số liệu chất lượng cho các chương trình BLS khác nhau và các nỗ lực định kỳ để phát triển các mô hình và công cụ trên toàn Cục để đánh giá chất lượng (ví dụ,

Dippo, 1997). Tuy nhiên, hiện tại không tồn tại khung đánh giá chất lượng toàn diện hoặc có hệ thống trong Cục.

Bài viết này tóm tắt một nỗ lực BLS OPLC để tạo ra một khung chất lượng cho bốn chương trình điều tra: Điều tra Chi tiêu của người tiêu dùng (CE), chỉ số giá tiêu dùng (CPI), chỉ số giá sản xuất (PPI) và Chương trình Giá quốc tế (IPP). Dự án này được chỉ đạo bởi Phó ủy viên của BLS OPLC, Michael Horrigan, trong nỗ lực củng cố, tối ưu hóa và ghi lại các quy trình chất lượng trong các chương trình này. Công việc được thúc đẩy bởi việc công nhận rằng mỗi lượng chương trình đã áp dụng nhiều chỉ tiêu chất

➤ ➤ ➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

cho cả quy trình nội bộ và cho đầu ra tương ứng của họ nhưng có sự không nhất quán về cả phạm vi bao phủ và chiều sâu của các chương trình. Trước khi phát triển một khung chung cho cả quy trình sản xuất nội bộ và về chất lượng đầu ra, OPLC đã liên hệ với Văn phòng Nghiên cứu Phương pháp điều tra (OSMR) của BLS và yêu cầu tóm tắt về các loại khung chất lượng tồn tại trên hệ thống thống kê của Hoa Kỳ, cộng đồng thống kê quốc tế và các tổ chức khu vực tư nhân tham gia thu thập và phổ biến dữ liệu. Sau khi xem xét mở rộng các quy trình còn tồn tại do các đồng tác giả OSMR cung cấp cho bài viết này, Scott Fricker và Polly Phipps, ban quản lý OPLC đã phát triển một khung đồng thuận về các biện pháp chất lượng để hướng dẫn đánh giá cả quy trình tác nghiệp điều tra nội bộ và các kết quả thống kê được tạo ra ở mỗi chương trình. Bài viết báo cáo về nỗ lực này. Phần 2 cung cấp một bản tóm tắt về đánh giá OSMR của tài liệu chất lượng. Phần 3 báo cáo về cuộc tranh luận của bốn chương trình OPLC để đạt được sự đồng thuận và Phần 4 tóm tắt khung đồng thuận và các số liệu chất lượng chính là kết quả của những cuộc tranh luận đó.

2. Khung chất lượng

BLS và các cơ quan thống kê liên bang khác của Hoa Kỳ cam kết mạnh mẽ để tạo ra số liệu thống kê chất lượng. Nhưng, ý nghĩa của “chất lượng” đối với chúng tôi là gì? Ở phần này chúng tôi xem xét các khái niệm khác nhau về chất lượng thống kê và chất lượng điều tra, xem xét hai khung chất lượng phổ biến được các tổ chức thống kê áp dụng. Chúng tôi cũng mô tả các tiêu chuẩn và hướng dẫn do Văn phòng Quản lý và Ngân sách Hoa Kỳ (OMB), cơ quan chịu trách nhiệm giám sát và đo lường chất lượng của các chương trình, chính sách và quy trình của cơ quan liên bang.

2.1 Đánh giá khung chất lượng

Có hai khung khái niệm chính liên quan đến chất lượng thống kê và chất lượng điều tra. Đầu tiên là khung tổng sai số điều tra (TSE), bắt nguồn từ tài liệu thống kê truyền thống và đặc biệt tập trung vào tính chính xác của dữ liệu điều tra. Khung thứ hai gồm nhiều yếu tố chất lượng, bao gồm sự phù hợp của một sản phẩm thống kê để sử dụng bởi khách hàng và người dùng và có nguồn gốc từ phong trào Quản lý chất lượng toàn diện (TQM) và các Khung quản lý chất lượng (QM) khác. Sự khác biệt chính giữa hai khung là ở thông số kỹ thuật của chúng về thông số chất lượng và tầm quan trọng lớn hơn vào nhu cầu của người dùng dữ liệu được tìm thấy trong các phương pháp QM. (Xem Lyberg, 2012 để biết cách xử lý triệt để hơn các nguyên tắc TSE và QM, và quan điểm lịch sử về sự phát triển của các khái niệm về chất lượng điều tra).

Tổng sai số điều tra là một khái niệm mô tả các thuộc tính sai số thống kê của ước tính điều tra bằng cách kết hợp tất cả các nguồn sai số có thể phát sinh trong quá trình điều tra. Khung tổng sai số điều tra tập trung vào chất lượng dữ liệu điều tra được đo bằng độ chính xác hoặc sai số bình phương trung bình (độ chệch và phương sai) của ước tính, với mục tiêu giảm các sai số điều tra quan trọng đối với chất lượng dữ liệu và giảm thiểu chi phí điều tra. Có một số khác biệt nhỏ trong cách xác định thuật ngữ “tổng sai số điều tra”, nhưng đã có sự thống nhất rộng rãi về các yếu tố cấu thành chính của nó. Ví dụ, các sai số thường được nhóm thành hai nhóm chính - sai số chọn mẫu và sai số phi chọn mẫu.

Trái ngược với các phương pháp tiếp cận chất lượng của TSE, các khung QM tập trung vào sự phù hợp của việc sử dụng các sản phẩm thống kê của các nhóm người dùng khác nhau khi xác định chất lượng và xác

định các biện pháp chất lượng. Loại khung này mở rộng khái niệm chất lượng thành nhiều chiều. Độ chính xác là thước đo chất lượng được xác định rõ ràng và định lượng nhất. Các thước đo khác mà người dùng có xu hướng ưu tiên bao gồm: Tính phù hợp, tính kịp thời, khả năng tiếp cận, khả năng giải thích và tính chặt chẽ. Tính phù hợp được định nghĩa là sản xuất thông tin về các khái niệm phù hợp và sử dụng các khái niệm đo lường phù hợp trong chủ đề. Việc thông tin có kịp thời và có thể truy cập được đối với người dùng hay không là hai yếu tố chất lượng bổ sung. Khả năng giải thích tập trung vào sự sẵn có của các khái niệm, các biến, phân loại, phương pháp thu thập, xử lý và ước tính cho người dùng để họ có thể tự đánh giá. Tính chặt chẽ liên quan đến thông tin phù hợp với các khung rộng lớn; việc sử dụng các khái niệm tiêu chuẩn, các biến và phương pháp phân loại; và nếu thông tin có thể được xác nhận với các bộ dữ liệu liên quan.

Các khung QM đã được áp dụng ở nhiều cơ quan thống kê quốc gia (ví dụ, xem Thống kê Canada, 2002, ví dụ ban đầu), mặc dù có sự khác biệt giữa các cơ quan và tổ chức. Ví dụ, tại Hoa Kỳ, Hội đồng Chính sách Thống kê Liên ngành đã đưa ra tính phù hợp, tính chính xác, tính kịp thời và phổ biến/khả năng tiếp cận (các biện pháp đầu ra thống kê), nhưng cũng bổ sung chi phí và thành tích nhiệm vụ như các tiêu chuẩn về hiệu suất của các cơ quan thống kê liên bang Hoa Kỳ. Khung quốc tế bao gồm chi phí cũng như gánh nặng của người trả lời như các quy trình thống kê và xem xét khả năng đo lường chi phí và gánh nặng quan trọng, trong đó cho phép đánh giá sự cân bằng của chi phí đối với lợi ích của dữ liệu chất lượng đầu ra.

Eurostat đã phát triển một khung chất lượng và Quy tắc thực hành rộng rãi dựa trên các tiêu chuẩn của Hệ thống Thống kê châu

Âu (ESS). Các tài liệu đưa ra các thông số kỹ thuật để đánh giá chất lượng và hiệu suất, bao gồm những gì cần được đưa vào báo cáo, các chỉ tiêu chất lượng sản phẩm cụ thể và đo lường các biến số chất lượng của quy trình. Ngoài ra, các cuộc điều tra người dùng, công cụ tự đánh giá, các công cụ kiểm toán, cũng như dán nhãn và chứng nhận được đề cập trong Cẩm nang Eurostat (Ehling và Korner, 2007). Quỹ Tiền tệ Quốc tế (2012) có khung đánh giá chất lượng dữ liệu được đặt ra cho các chỉ tiêu chính, bao gồm CPI và PPI. Các chiều bao gồm: Tính toàn vẹn, tính đúng đắn về phương pháp, độ chính xác và độ tin cậy, khả năng phục vụ và khả năng tiếp cận. Các chiều được tập trung vào chỉ tiêu là thước đo thống kê, nhưng đã được kết nối với khung ESS (Laliberte, Grunewald và Probst, 2004).

Nhiều cơ quan đã đi sâu vào chi tiết để xác định các chỉ tiêu và các mặt hàng đo lường chất lượng. Thông thường, các bước của quy trình điều tra được sử dụng như một phần của khung và các bước điều tra có thể rất phổ quát hoặc chi tiết. Với mức độ chi tiết mà các cơ quan đã đặt ra, việc ưu tiên các biện pháp chất lượng thường hữu ích; ví dụ, Văn phòng Thống kê Quốc gia (ONS) ở Anh xác định một danh sách ngăn các biện pháp chất lượng sản phẩm chính. Thống kê Canada cung cấp các hướng dẫn và chỉ tiêu chất lượng cho mỗi 17 bước trong một cuộc điều tra (2009). ONS tập trung vào 8 loại chính (2013). Ngoài ra, Eurostat đã đưa ra một số lượng nhỏ các biện pháp chất lượng cụ thể cho các chỉ tiêu kinh tế, bao gồm chỉ số hài hòa giá tiêu dùng và chỉ số sản phẩm công nghiệp (Mazzi và cộng sự, 2005). Nó cũng đã ban hành các hướng dẫn để cải thiện chất lượng của các quy trình sản xuất điều tra, thông qua việc xây dựng các biểu đồ luồng quy trình, xác định và giám sát các

➤ ➤ ➤ THỐNG KÊ QUỐC TẾ VÀ HỘI NHẬP

biến quy trình quan trọng trong các giai đoạn khác nhau của chu kỳ điều tra¹. Những hướng dẫn này mở rộng công việc trước đó của Biemer và Caspar (1994) và Morganstein và Marker (1997), đồng thời cung cấp các kỹ thuật để phân tích các biến trong quy trình vận hành theo thời gian để cải thiện khả năng của các quy trình đó.

2.2 Chương trình Chất lượng trong các cơ quan thống kê khác của Hoa Kỳ

Hệ thống thống kê của Hoa Kỳ là một mạng lưới các cơ quan thống kê phi tập trung, nhưng nó hoạt động theo các tiêu chuẩn chất lượng do OMB đặt ra. Năm 2002, OMB đã ban hành các hướng dẫn cuối cùng để đảm bảo và tối đa hóa chất lượng thông tin được phổ biến bởi các cơ quan liên bang². Thuật ngữ chất lượng được sử dụng bao gồm tiện ích, tính khách quan và tính toàn vẹn. Tiện ích đề cập đến tính hữu ích của thông tin cho người dùng. Tính khách quan bao gồm liệu thông tin phổ biến được trình bày một cách chính xác, rõ ràng, đầy đủ và không thiên vị, và vấn đề căn bản là chính xác. Tính toàn vẹn liên quan đến bảo mật hoặc bảo vệ thông tin khỏi sự truy cập hoặc sửa đổi trái phép, hoặc thỏa hiệp thông tin thông qua tham nhũng hoặc giả mạo. Các cơ quan đã phát triển các hướng dẫn chất lượng thông tin của riêng họ; các tài liệu hướng dẫn kết quả từ các cơ quan không bao gồm bất kỳ loại biện pháp thực hiện. BLS đã giải quyết các hướng dẫn trên một trang web với thảo luận về các hướng dẫn khác nhau và cách chúng được đáp ứng, cũng như một

phần với các hướng dẫn cụ thể về tính toàn vẹn dữ liệu. Phần toàn vẹn dữ liệu thảo luận về quy trình bảo mật, quy trình an toàn và an ninh, thu thập và phổ biến dữ liệu³. Tất cả các cơ quan thống kê liên bang khác cũng phát triển hướng dẫn.

Năm 2006, OMB đưa ra 20 tiêu chuẩn và hướng dẫn cho các điều tra thống kê bao gồm quy trình điều tra: Phát triển các khái niệm, phương pháp và thiết kế; thu thập dữ liệu; xử lý và hiệu chỉnh dữ liệu, đưa ra các ước tính và dự đoán, phân tích dữ liệu, xem xét quy trình và phổ biến các sản phẩm thông tin⁴. Các biện pháp hiệu suất cụ thể liên quan đến các nguyên tắc này bao gồm phân tích độ chệch không trả lời khi đơn vị không trả lời dưới 80% hoặc mục trả lời dưới 70% đối với bất kỳ hạng mục nào được sử dụng trong báo cáo hoặc nghiên cứu độ chệch phạm vi khi tỷ lệ phạm vi giảm xuống dưới 85%.

Cục Điều tra Dân số Hoa Kỳ là cơ quan duy nhất chúng tôi xác định là đã phát triển tài liệu về các tiêu chuẩn này để cung cấp hướng dẫn bổ sung về các chương trình, hoạt động và bao gồm các vấn đề về phương pháp và hoạt động độc đáo⁵. Chúng căn chỉnh các tiêu chuẩn với các tiện ích, tính khách quan và kích thước toàn vẹn được đặt ra bởi OMB. Tài liệu bao gồm các hướng dẫn chi tiết và kỹ lưỡng, định nghĩa và các yêu cầu cho tất cả các hoạt động, kỹ thuật, quy trình và hệ thống liên quan đến từng giai đoạn của quy trình điều tra. Nói chung, tài liệu không bao gồm các biện pháp thực hiện.

¹ <http://epp.eurostat.ec.europa.eu/portal/page/portal/auality/documents/HANDBOOK%20ON%20IMPROVING%20QUALITY.pdf>

² http://www.whitehouse.gov/omb/fedreg_final_information_quality_guidelines

³ <http://www.bls.gov/bls/quality.htm>

http://www.bls.gov/bls/data_integrity.htm

⁴ http://www.whitehouse.gov/sites/default/files/omb/infocreg/statpolicy/standards_stat_surveys.pdf

⁵ <http://www.census.gov/auality/standards/index.html>

3. Các ý kiến về Khung chất lượng cho OPLC

Được hưởng lợi từ đánh giá tài liệu và các cuộc họp giao ban do OSMR cung cấp, ban quản lý OPLC đặt ra về nhiệm vụ phát triển khung đồng thuận. Bước đầu tiên là phát triển một tài khoản chi tiết về các biện pháp chất lượng được áp dụng bởi mỗi chương trình để xác định các lĩnh vực phổ biến và khác biệt. Điều này đã được thực hiện thông qua một loạt các bản ghi nhớ và tóm tắt chi tiết được trao cho ban quản lý OPLC về các biện pháp chất lượng được áp dụng bởi mỗi chương trình.

Như các cuộc họp giao ban đã chỉ ra, một số biện pháp chất lượng có điểm chung và một số có sự khác biệt giữa bốn chương trình OPLC. Đồng thời, có một số biện pháp rất độc đáo trong ba chương trình giá do bản chất của thống kê giá. Ví dụ, như được mô tả tiếp theo, có các biện pháp kiểm tra các tùy chọn được thực hiện khi hàng hóa hoặc dịch vụ được “định giá lại” tại một cơ sở đã thay đổi về chất lượng hoặc không còn có sẵn để định giá. Một mẫu xe mới, các dòng quần áo đã ngừng sản xuất và các mặt hàng nhập khẩu thay đổi nước xuất xứ chỉ là ba ví dụ về những thách thức độc đáo mà các chương trình giá phải đối mặt trong nỗ lực đẩy lùi hàng hóa hoặc dịch vụ chính xác từ tháng này sang tháng khác. Ngoài ra, chương trình thứ tư, điều tra CE, là điều tra hộ gia đình, giới thiệu một số yếu tố độc đáo liên quan đến các mục và điều tra không trả lời mà không có trong thiết lập dựa trên các chương trình giá cơ sở.

Một trong những lợi ích bất ngờ của việc chia sẻ thông tin này là việc xác định các thông lệ thu thập dữ liệu theo từng cuộc điều tra được các chương trình khác quan tâm. Không nơi nào xác thực hơn so với bài thuyết trình của CPI về giám sát thay thế của họ.

Một trong những tính năng quan trọng của việc thu thập dữ liệu giá là các tùy chọn được thực hiện khi mặt hàng trong rổ hàng hóa được chọn để định giá mỗi tháng đã thay đổi về nội dung hoặc không còn có sẵn để định giá. Các tùy chọn được chọn bởi người thu thập dữ liệu được điều chỉnh bởi các quy tắc và quy trình thu thập dữ liệu chi tiết và có thể liên quan đến việc xác định rằng thay đổi trong nội dung của mặt hàng không phải là xác định giá (chẳng hạn như thay đổi màu sắc), ước tính giá trị của thay đổi chất lượng trong trường hợp thay đổi là đáng kể (và quyết định sử dụng phương pháp thay đổi chất lượng nào phù hợp nhất), thay thế cho một mặt hàng tương tự khác (và điều chỉnh chất lượng giữa hai mặt hàng), quyết định rằng mặt hàng đó tạm thời không có (và sử dụng phương pháp gán giá được ghi chép trước đó để chuyển tiếp), hoặc ngừng mặt hàng và bắt đầu một loạt giá mới cho một mặt hàng mới. CPI ghi lại mọi tùy chọn được thực hiện cho mọi mặt hàng và kiểm tra xu hướng trong các tùy chọn này theo loại mặt hàng. Các chương trình PPI và IPP kể từ đó đã áp dụng các phiên bản của việc theo dõi thay thế này.

Một khía cạnh bất ngờ khác của việc chia sẻ các phương pháp đo lường chất lượng này, một khía cạnh quan trọng trong khung chất lượng đồng thuận cuối cùng của chúng tôi, đó là sự phân chia rõ ràng tồn tại giữa PPI và IPP, nói cách khác với tầm quan trọng tương đối lớn hơn đối với đo lường quy trình tác nghiệp và CPI, và các biện pháp liên quan đến đo lường đầu ra. Tất cả ba điều tra đều có cả hai khía cạnh, đó là tầm quan trọng tương đối và số lượng các biện pháp trong từng lĩnh vực là bất ngờ. IPP tạo ra một báo cáo cực kỳ chi tiết hàng tháng về các biện pháp xử lý chi phối mọi khía cạnh của quy trình tác nghiệp. Như một ví dụ minh họa, một trong những biện pháp chất lượng là

➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

“100% của tất cả TSUSA/B ánh xạ sang SCG và TSUSA/B đến ánh xạ thứ cấp là chính xác trước khi chạy các chỉnh sửa khung trong tuần thứ ba của tháng Hai và tháng Chín mỗi năm”. PPI có một tầm quan trọng tương tự về các biện pháp xử lý, dấu không hoàn toàn chi tiết. Mặc dù CPI có một số biện pháp xử lý, nhưng chúng cũng không chi tiết như các biện pháp được theo dõi bởi IPP hoặc PPI. Khi xây dựng khung đồng thuận, một quyết định đã được đưa ra bao gồm một phần riêng về các biện pháp quy trình tác nghiệp, như sẽ thấy trong phần tiếp theo, chúng tôi đã giải quyết một tập hợp các biện pháp có tính phổ biến cao hơn trong ba chương trình giá, trong khi thừa nhận rằng một số chương trình, đặc biệt là IPP, có thể chọn theo dõi các biện pháp xử lý ở mức độ chi tiết tốt hơn nhiều so với các chương trình khác.

Trái ngược với IPP và PPI, CPI cũng có số lượng đo lường tương đối lớn hơn liên quan đến chất lượng đầu ra. Một lý do quan trọng cho sự khác biệt này giữa CPI và PPI/IPP là việc trước đây tạo ra các ước tính sai số chuẩn hàng năm cho từng chỉ số, trong khi đó, cả PPI và IPP đều không xây dựng bất kỳ biện pháp nào khác nhau cho các chỉ số của mình. Thật vậy, một phần của con đường đạt được khung đồng thuận là công việc rất tích cực mà cả PPI và IPP đang thực hiện để xây dựng những phương sai như vậy trong tương lai và các loại đo lường chất lượng liên quan đến sai số điều tra từ CPI rất đáng quan tâm.

Điều tra CE, như điều tra hộ gia đình, khác biệt với ba chương trình giá ở nhiều khía cạnh. Một khía cạnh đã trở nên rõ ràng ngay lập tức, đó là thực tế rằng nguyên tắc hướng dẫn CE trong việc phát triển khung chất lượng là việc áp dụng rõ ràng phương pháp đo lường sai số điều tra tổng thể. Có một số yếu tố góp phần vào việc thông qua

của mục tiêu của CE để giảm thiểu tổng sai số điều tra. Điểm mấu chốt trong số này là việc đánh giá thấp chi tiêu của người tiêu dùng đối với một số mặt hàng nhất định (đặc biệt là các mặt hàng liên quan đến thói quen cá nhân như chi tiêu cho thuốc lá hoặc rượu) được biết là không phù hợp với ước tính có được từ dữ liệu bán hàng kinh doanh độc lập. Đồng thời, CE đang thực hiện thiết kế lại điều tra chính và việc giảm sai số điều tra có thể kiểm chứng là một nguyên tắc chỉ đạo của nỗ lực thiết kế lại (cùng với việc giảm gánh nặng tương ứng). Tầm quan trọng của sai số điều tra và mối quan tâm về việc đánh giá thấp đã khiến CE quan tâm đến các khía cạnh cụ thể của chất lượng không có cùng tầm quan trọng hoặc ưu tiên trong ba chương trình giá. Điều này đặc biệt đúng trong những nỗ lực của CE để kiểm tra sự gắn kết của các ước tính của họ so với các điều tra và nguồn dữ liệu hành chính khác, cả về mức độ chi tiêu của người tiêu dùng và phân phối chi tiêu theo loại mặt hàng.

Như với bất kỳ thước đo chất lượng nào, việc đưa ra kết luận như vậy về tầm quan trọng tương đối của các biện pháp chất lượng nhất định so với các biện pháp khác thường là vấn đề mức độ. Trong các chương trình giá, không có sự tồn tại của các biện pháp thay thế đáng kể của giá sản xuất hoặc tiêu dùng, người ta chú ý nhiều hơn đến hành vi hoặc chỉ số giá liên quan đến các chỉ số giá khác (Ví dụ: Hành vi giá biến động của đầu vào đến sản xuất có được phản ánh một phần trong sự biến động của các sản phẩm mà đầu vào có giá trị đáng kể so với tổng chi phí sản xuất) hoặc hành vi của các chỉ số giá theo thời gian. Điều đang được nói đến, CPI gần đây đã bắt tay vào một loạt các dự án thú vị để so sánh việc phân phối các mặt hàng được chọn để chọn mẫu với các phân phối ngụ ý bởi dữ liệu máy quét cho cùng

một mặt hàng, kiểm tra sự gắn kết của quy trình chọn mẫu nội bộ.

4. Hướng tới Khung chất lượng chung

Để đi đến một khung đồng thuận, chúng tôi đã dựa vào việc yêu cầu các chương trình để phản ứng với các phiên bản nhỏ của khung chất lượng bao gồm các biện pháp phù hợp với những gì chúng tôi học được từ cả đánh giá tài liệu và chia sẻ thực tiễn trên các chương trình (và sau đó thảo luận). Chúng tôi đã sớm quyết định phát triển một hệ thống gồm hai phần, một phần nhấn mạnh vào chất lượng đầu ra và phần thứ hai nhấn mạnh vào các quy trình tác nghiệp làm phát sinh các đầu ra này. Các đo lường chất lượng đầu ra được chia thành 14 loại riêng biệt: (1) Mức độ liên quan; (2) Độ chính xác và độ tin cậy - sai số chọn mẫu; (3) Độ chính xác và độ tin cậy - sai số phi chọn mẫu; (4) Phạm vi sai số; (5) Sai số đo lường; (6) Sai số xử lý; (7) Sai số hiệu chỉnh; (8) Sai số mô hình hóa; (9) Tính kịp thời; (10) Khả năng tiếp cận; (11) Khả năng giải thích; (12) Tính chặt chẽ; (13) Chi phí; và (14) Uy tín, liêm chính và bảo mật.

Khung quy trình tác nghiệp dựa trên các quy trình tác nghiệp chính có liên quan đến cả ba chương trình giá và CE: (1) Thiết kế/khái niệm; (2) Chọn mẫu; (3) Thu thập dữ liệu; (4) Ước tính; (5) Cập nhật trọng số; và (6) Phổ biến.

Trong nhiều khía cạnh, đặc điểm kỹ thuật của các biện pháp đầu ra chất lượng ngay lập tức dễ tiếp cận hơn nhờ sự nhấn mạnh tương đối vào chất lượng trong tài liệu học thuật và tính phổ biến lớn hơn giữa các chương trình hoặc trong các biện pháp mà họ áp dụng, hoặc trong trường hợp phương sai, các biện pháp họ muốn thông qua trong tương lai. Như đã đề cập ở trên, các biện pháp quy trình tác nghiệp rất khác nhau về mức độ chi tiết tương đối của chúng, do đó,

sự đồng thuận mà bạn sẽ thấy dưới đây nhằm tìm ra nhiều điểm chung nhất có thể, cho phép tồn tại các biện pháp chi tiết cụ thể hơn.

Khi chúng tôi bắt đầu chỉ định các biện pháp chất lượng của mình, một khía cạnh của đánh giá tài liệu Fricker/Phipps trở nên rõ ràng hơn là cần phải phát triển các biện pháp liên quan đến nhu cầu của người dùng, các biện pháp không được đề cập rõ ràng trong bất kỳ biện pháp chất lượng cụ thể nào của chương trình. Điều này không có nghĩa là các chương trình riêng lẻ không cố gắng đo lường mức độ đáp ứng nhu cầu của người dùng (mặc dù các thực tiễn khác nhau), nhiều hơn là họ đã không đưa các thực tiễn đó vào các khung chương trình của riêng họ. Điểm thứ hai này đặc biệt trong bảng chúng với các biện pháp kịp thời và khả năng tiếp cận.

Cuối cùng, như hai trình bày được đưa ra dưới đây cho thấy, có một số sự chùng chèo tự nhiên giữa sản lượng cụ thể và các biện pháp quy trình tác nghiệp mà chúng tôi đã áp dụng. Chúng tôi đã có một số cuộc thảo luận cố gắng phát triển các biện pháp loại trừ lẫn nhau qua hai mô hình nhưng cuối cùng chúng tôi quyết định cho phép sao chép như vậy vì lý do đơn giản là mỗi mô hình có các biện pháp chất lượng duy nhất không tìm thấy trong mô hình khác cũng như có các biện pháp phổ biến. Các phương pháp này đã được tìm thấy cho mỗi phương pháp có giá trị gia tăng về mặt quản lý các chương trình của chúng tôi và đưa ra một quan điểm khác nhau về bản chất của chất lượng trong các quy trình và đầu ra điều tra của chúng tôi.

Bảng 1 hiển thị các biện pháp chất lượng đầu ra từ khung đồng thuận OPLC; Bảng 2 trình bày các biện pháp chất lượng quy trình tác nghiệp. Các bảng này liệt kê các kích thước chất lượng chính, định nghĩa của chúng và các số liệu chất lượng liên quan cho cả chất lượng sản phẩm và quy trình.

Bảng 1: Các biện pháp chất lượng đầu ra của OPLC

Thông số chất lượng	Định nghĩa	Các biện pháp chất lượng đầu ra của OPLC
1. Tính phù hợp	Dữ liệu có thỏa mãn nhu cầu người dùng không? Là chương trình sản xuất thông tin về các khái niệm đúng hay không?	<p>Chương trình cung cấp một tuyên bố tường thuật về sự liên quan của dữ liệu của họ. Điều này bao gồm mô tả mục tiêu đo lường của chương trình, dữ liệu được tạo ra, hệ thống phân loại được sử dụng và phạm vi bảo hiểm của các cập nhật sản phẩm dữ liệu.</p> <p>Chương trình thu thập phản hồi từ người dùng dữ liệu về sự hài lòng và nhu cầu dữ liệu của họ thông qua nhiều cơ chế phản hồi một cách thường xuyên.</p> <p>Chương trình có một mô tả bằng văn bản xác định các khoảng trống đã biết trong việc đáp ứng nhu cầu và ưu tiên của người dùng để khép lại các khoảng trống đó. Tài liệu này được cập nhật một cách thường xuyên.</p>
2a. Tính chính xác và độ tin cậy - sai số chọn mẫu	Là chương trình giảm thiểu tổng sai số điều tra ngẫu nhiên?	<p>Sai số chuẩn Hệ số biến thiên Sai số chuẩn tương đối Khoảng tin cậy</p>
2b. Sai số phạm vi	Sai số phạm vi phát sinh do thiếu sót, trùng lặp và sai số nội dung. Chương trình có cố gắng đo lường sát giữa tổng thể phạm vi và tổng thể mục tiêu không?	<p>Tỷ lệ ngoài phạm vi Tỷ lệ phân loại sai Dưới và ngoài phạm vi</p>
2c. Sai số không trả lời	Sai số không trả lời do không thể thu thập thông tin đầy đủ về tất cả các đơn vị trong một mẫu được chọn. Chương trình có tính toán cả tỷ lệ đơn vị và mục trả lời không?	<p>Tỷ lệ đơn vị trả lời (có trọng số và không có trọng số) Tỷ lệ trả lời các mục chính % trọng số cuối cùng từ các ô không được gán cho các hạng mục ước tính chính Tỷ lệ xung đột</p>
2d. Sai số đo lường	Sai số đo lường xảy ra khi đáp ứng được cung cấp khác với giá trị thực và có thể được quy cho người trả lời, người phỏng vấn, bảng câu hỏi hoặc phương pháp thu thập. Chương trình có cố gắng so sánh kết quả với các nguồn khác không?	<p>So sánh điểm chuẩn Quy trình thay thế Phạm vi của điều này bao gồm: - Tỷ lệ thay thế - Phần trăm các mục được liên kết - Phần trăm chất lượng mặt hàng được điều chỉnh - Phần trăm hạng mục không tìm thấy vật thay thế</p>

Thông số chất lượng	Định nghĩa	Các biện pháp chất lượng đầu ra của OPLC
2e. Sai số xử lý	Sai số xử lý là sai số trong kết quả điều tra cuối cùng phát sinh từ việc thực hiện sai các phương pháp thực hiện được lên kế hoạch chính xác. Ví dụ, chương trình có theo dõi số lượng lỗi mã hóa hoặc nhập dữ liệu được phát hiện do kết quả của một cuộc phỏng vấn lại không?	Tỷ lệ sai số chỉnh sửa Tỷ lệ sai số mã hóa Tỷ lệ sai số nhập dữ liệu
2f. Sai số hiệu chỉnh	Chương trình có theo dõi số lượng và kích thước tương đối của các phiên bản giữa xuất bản ban đầu và cuối cùng không?	Kích thước tương đối của các sửa đổi trong các khái niệm ước tính chính giữa xuất bản đầu tiên và cuối cùng
2g. Sai số mô hình hóa	Sai số mô hình hóa xảy ra khi sử dụng các phương pháp, như hiệu chỉnh, hồi quy tổng quát, điều chỉnh theo mùa vụ và các mô hình khác không có trong các thành phần chính xác trước đó, để tính toán thống kê hoặc chỉ số. Chương trình có cố gắng để tìm hiểu ảnh hưởng của các sai số liên quan đến mô hình của chúng không?	Mức độ phù hợp của số liệu thống kê phù hợp cho loạt điều chỉnh theo mùa vụ
3. Tính kịp thời	Các ước tính điều tra được báo cáo kịp thời để tối đa hóa tính hữu ích của chúng? Là dữ liệu và tài liệu liên quan được phát hành theo lịch trình?	Chương trình cung cấp một bài tường thuật mô tả các dòng sản phẩm của họ, tính định kỳ của chúng và đánh giá tổng thể về tính kịp thời. Số lượng tin tức được phát hành đúng giờ mỗi năm, theo một lịch trình được công bố trước Độ trễ giữa kết thúc thu thập dữ liệu và công bố ước tính
4. Khả năng tiếp cận	Chương trình có cung cấp các công cụ trích xuất dữ liệu thân thiện với người dùng không? Là truy cập vào dữ liệu giá cả phải chăng?	Chương trình cung cấp một tường thuật về khả năng tiếp cận dữ liệu và sản phẩm của chúng. Đánh giá của người dùng về khả năng truy cập Số lượng người đăng ký vào các ấn phẩm chương trình Số lượng yêu cầu thông tin được ghi nhận hàng tháng Số lần truy cập web tổng thể và theo sản phẩm, hàng tháng
5. Khả năng giải thích	Chất lượng và phạm vi của tài liệu điều tra giải thích các khái niệm và phương pháp điều tra là gì?	Chương trình cung cấp một tường thuật mô tả tài liệu của họ có sẵn trên các khái niệm và phương pháp luận Đánh giá của người dùng về khả năng diễn giải

➤ ➤ ➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

Thông số chất lượng	Định nghĩa	Các biện pháp chất lượng đầu ra của OPLC
6. Tính chặt chẽ	Mức độ mà dữ liệu được lấy từ các nguồn khác nhau nhưng đo lường cùng một hiện tượng, tương tự như các ước tính được tạo ra bởi chương trình?	Chương trình tiến hành các nghiên cứu đang thực hiện so sánh các ước tính chính của nó với các ước tính có nguồn gốc từ các nguồn thay thế.
7. Giá cả	Chi phí thu thập, xử lý và phổ biến dữ liệu là gì?	Chi phí được nhận trực tiếp và đầy đủ cho mỗi đơn vị được bắt đầu (PPI/IPP/CPI C & S/CPI nhà ở/Nhật ký CE/Phòng vấn CE) được phân tầng theo phương thức thu thập Chi phí được nhận trực tiếp và đầy đủ cho mỗi đơn vị được phỏng vấn lại (PPI/IPP/CPI C & S/CPI nhà ở/Nhật ký CE/Phòng vấn CE) được phân tầng theo phương thức thu thập Số lượng và phân phối người trả lời theo phương pháp thu thập dữ liệu (PPI/IPP/Nhật ký CE/Phòng vấn CE)
8. Tính tin cậy, liêm chính, bảo mật	Các ước tính được tạo ra bởi chương trình điều tra được xem là đáng tin cậy đến mức độ nào? Ở mức độ nào thì chương trình và tổ chức cơ bản được xem là có tính toàn vẹn? Chương trình có bảo vệ tính bảo mật của người trả lời không?	Vi phạm dữ liệu được báo cáo mỗi quý hoặc năm Số lỗi được tìm thấy trong các số được công bố sau khi phát hành

Bảng 2: Đo lường chất lượng quy trình tác nghiệp

Thông số chất lượng	Định nghĩa	Các biện pháp chất lượng đầu ra của OPLC
1. Thiết kế/Khái niệm	Quy trình tác nghiệp này xác định các mục tiêu ước tính chính của chương trình điều tra và cách các mục tiêu này được đo thông qua các công cụ và quy trình thu thập dữ liệu.	Chương trình cung cấp một tường thuật về (các) mục tiêu ước tính của chương trình điều tra và cơ sở lý thuyết/khái niệm cho mục tiêu đó.
2. Chọn mẫu	Quy trình tác nghiệp này xác định (các) khung chung được sử dụng để chọn mẫu để thu thập dữ liệu, quy trình chọn mẫu và các quy trình và định kỳ cho việc lấy lại mẫu (nếu phù hợp) và phương pháp điều chỉnh mức độ mất mẫu theo thời gian.	Chương trình cung cấp một tuyên bố tường thuật mô tả các khung chung có liên quan, quá trình chọn mẫu và các phương pháp để điều chỉnh sự mất mẫu theo thời gian Tuổi trung bình của mẫu tổng thể và theo các đặc điểm có liên quan như ngành và sản phẩm, nhân khẩu học... <u>Chỉ có IPP/PPI</u> So sánh kích thước mẫu được nhằm mục tiêu so với thực tế

Thông số chất lượng	Định nghĩa	Các biện pháp chất lượng đầu ra của OPLC
3. Thu thập dữ liệu	<p>CPI, PPI, IPP</p> <p>Trong Chương trình giá BLS, quy trình tác nghiệp này mô tả bộ sưu tập giá đầu tiên được thu thập tại một cơ sở, bao gồm các bước phân chia được thực hiện bởi Nhà kinh tế thực địa để có được hàng hóa và dịch vụ chính xác được định giá và tính toán lại các mặt hàng đó.</p> <p>Trong chương trình CE, quy trình tác nghiệp này bao gồm việc thu thập dữ liệu cho 5 cuộc phỏng vấn hàng quý trong Điều tra hộ gia đình CE và thu thập dữ liệu cho Nhật ký CE 2 tuần.</p>	<p>Tất cả</p> <p>Tỷ lệ đơn vị trả lời (có trọng số và không trọng số)</p> <p>Ước tính độ dài trung bình của một cuộc phỏng vấn bằng phương pháp thu thập dữ liệu</p> <p>Số lượng và phân phối người trả lời theo phương pháp thu thập dữ liệu</p> <p>Tỷ lệ đơn vị trả lời cụ thể của CPI/PPI, IPP (có trọng số và không có trọng số) để bắt đầu và đánh giá lại</p> <p>Tỷ lệ thay thế</p> <p>Tỷ lệ điều chỉnh chất lượng theo phương pháp</p>
4. Dự toán	<p>Quy trình tác nghiệp này bao gồm nhiều bước cần thiết để dịch dữ liệu được thu thập thông qua việc thu thập dữ liệu và phát triển các ước tính để xuất bản. Các bước này bao gồm chỉnh sửa, cắt bỏ, điều chỉnh trọng lượng, điều chỉnh theo mùa vụ, phân tích ngoại lệ và chuẩn bị các tệp ước tính và xem xét và đăng xuất cuối cùng</p>	<p>Chương trình cung cấp một tường thuật mô tả phương pháp ước tính của nó</p> <p>Tỷ lệ tranh chấp</p> <p>% trọng số cuối cùng từ các ô không bị tranh chấp cho các hạng mục ước tính chính</p> <p>Chương trình tiến hành các nghiên cứu đang thực hiện so sánh các ước tính chính của nó với các ước tính có nguồn gốc từ các nguồn thay thế.</p> <p><u>Chỉ có tỷ lệ thay thế của IPP/PPI</u></p> <p>Kích thước tương đối của các sửa đổi trong các khái niệm ước tính chính giữa xuất bản đầu tiên và cuối cùng</p>
5. Cập nhật trọng số	<p>Quy trình tác nghiệp này mô tả sự kết hợp của các trọng số chung mới đã được cập nhật (chẳng hạn như thông qua chính quyền của Cục Tổng Điều tra Dân số thập niên mới hoặc Điều tra dân số mới của Quinquennial, giữa những yếu tố khác).</p>	<p>Chương trình cung cấp một bài tường thuật mô tả kế hoạch trọng số của nó và cách cập nhật trọng số</p>
6. Tính phổ biến	<p>Quy trình tác nghiệp này mô tả việc đăng thông cáo báo chí điều tra, cơ sở dữ liệu, phân tích dữ liệu, tài liệu phương pháp điều tra, trong số những cách khác và cách cộng đồng người dùng dữ liệu có thể truy cập và sử dụng các tài liệu được xuất bản của chúng tôi.</p>	<p>Chương trình cung cấp một tường thuật về các sản phẩm và dịch vụ mà nó cung cấp cho công chúng.</p> <p>Đánh giá của người dùng về hiệu quả của việc phổ biến</p> <p>Tóm tắt về hiệu suất đúng giờ để phát hành chuỗi dữ liệu và tệp dữ liệu vi mô</p> <p># của loạt dữ liệu được công bố</p> <p>Số lần truy cập trên web tổng thể và theo sản phẩm, hàng tháng</p>

5. Thảo luận

Khung chất lượng được phát triển cho các chương trình giá BLS dựa trên cách tiếp cận của Eurostat, Thống kê Thụy Điển và các cơ quan thống kê khác ngoài Hoa Kỳ để phát triển các chỉ số chất lượng cho các điều tra của họ. Họ áp dụng một mô hình chất lượng đa chiều phù hợp với các phương pháp dựa trên QM và xác định các biện pháp chất lượng cho cả đầu ra sản phẩm và quy trình tác nghiệp điều tra. Mục tiêu của khung này là: (1) Xác định (tiềm năng) nguồn sai số; (2) Phát triển nhận thức về các rủi ro tương đối của các sai số này và định lượng chúng nếu có thể; (3) Xác định các lỗ hổng trong phương pháp luận; (4) Thúc đẩy việc sử dụng và tích hợp thông tin giám sát; và (5) Giúp ưu tiên các lĩnh vực cần cải thiện trong hoạt động điều tra.

Khi phát triển khung này, OPLC đã kết hợp nhiều tính năng được tìm thấy trong các phương pháp tiếp cận để đánh giá chất lượng điều tra và tận dụng chuyên môn của đội ngũ nhân viên của họ, những người quản lý và thực hiện các hoạt động điều tra cụ thể. Chúng tôi cung cấp một số quan sát về quá trình phát triển và khung kết quả. Đầu tiên, sự đa dạng của các chương trình OPLC - bao gồm cả điều tra hộ gia đình và cơ sở cá thể, các loại quy trình tác nghiệp khác nhau và các biện pháp chất lượng dành riêng cho chương trình khác nhau về số lượng và mức độ chi tiết - khiến cho việc tạo ra một khung toàn diện, duy nhất gặp khó khăn. Các tình hình lặp đi lặp lại là cần thiết để xây dựng tính linh hoạt đầy đủ thành cách tiếp cận đồng thuận phù hợp giữa các chương trình và sẽ cần thêm công việc để vận hành một bộ đầy đủ các biện pháp chất lượng và phát triển các hướng

dẫn cụ thể theo chương trình cho các số liệu đó. Thứ hai, như đã lưu ý ở trên, có sự chồng chéo đáng kể giữa các biện pháp chất lượng sản phẩm và quy trình được liệt kê trong Bảng 1 và Bảng 2, và tương đối ít các biện pháp xử lý có sẵn trong thời gian thực (hoặc gần thời gian thực), giới hạn các hoạt động giám sát và hoạt động can thiệp một liên kết với các phương pháp kiểm soát quá trình thống kê điển hình.

Bài viết trình bày về khung đồng thuận, trong đó nhấn mạnh các yếu tố và thước đo chất lượng phổ biến trong bốn chương trình, tuy nhiên, che giấu dữ liệu và quy trình chi tiết hơn đang được sử dụng để đánh giá chất lượng trong hai chương trình OPLC. Ví dụ: Chương trình IPP thường xuyên xuất bản các báo cáo chất lượng nội bộ theo dõi tiến trình trên nhiều mục tiêu chất lượng quy trình (ví dụ: Phần trăm các mặt hàng được trả lại, phần trăm của các tầng được xuất bản, tính kịp thời của các sửa đổi và ấn phẩm, ...). Tương tự, chương trình CE đã phát triển một bộ số liệu để đánh giá tác động của thay đổi thiết kế đối với chất lượng sản phẩm và quy trình (dựa trên thông tin lịch sử liên hệ và ước tính điều tra cho các đơn vị mẫu) và đề xuất các quy trình mới để phát triển, thực hiện và báo cáo các số liệu đó. Do đó, một quan sát thứ ba là các chương trình và cơ sở hạ tầng đánh giá chất lượng hiện có của chương trình sẽ ảnh hưởng đến tốc độ có thể tích hợp và nâng cao khung chất lượng tổng thể được trình bày ở đây. Việc triển khai rộng hơn khuôn khổ này trong OPLC sẽ là một quá trình đang diễn ra khi các chương trình của nó tiếp tục phát triển các sáng kiến chất lượng.

(Xem tiếp trang 13)

Tiếp theo trang 32

Chúng tôi cũng hy vọng rằng các chương trình BLS khác sẽ áp dụng và giúp tinh chỉnh khung chất lượng này. Mặc dù có những thách thức đối với nỗ lực đó (ví dụ: Phân cấp tổ chức, sự phức tạp gia tăng của việc đáp ứng nhu cầu chương trình, nguồn lực và cân nhắc hoạt động đa dạng), cũng có những cơ hội thực sự để thúc đẩy các hoạt động đánh giá chất lượng hiện có trong mỗi chương trình để xây dựng một hệ thống linh hoạt nhưng có hệ thống khung để quản lý chất lượng trong các sản phẩm và hoạt động điều tra BLS.

Hoàng Linh (dịch)

Nguồn:

<https://www.bls.gov/osmr/pdf/st140050.pdf>