

# Thực hiện một nghiên cứu định lượng (Conducting an empirical study)

Lê Việt Phú  
Trường Chính sách Công và Quản lý Fulbright

Ngày 9 tháng 5 năm 2018

# Table of contents

Phần I: Thực hiện một nghiên cứu định lượng

→ Cấu trúc một bài nghiên cứu định lượng

→ Một số vấn đề đáng lưu ý

→ Đánh giá một bài nghiên cứu định lượng

Phần II: Kỹ thuật và mô hình sử dụng trong nghiên cứu định lượng

# Thực hiện một nghiên cứu định lượng

Bản chất một bài nghiên cứu (thực nghiệm):

- ▶ Nghiên cứu là một quá trình khám phá ra cái mới: có thể là những quan hệ thực nghiệm mới, những lý thuyết mới chuyên sâu, hay hiểu biết sâu hơn về một hoặc một số vấn đề. Do đó bài nghiên cứu phải có *tính mới*, thay vì thu thập lại và diễn giải lại những kiến thức đã có.
- ▶ Nhưng mục đích của nghiên cứu không phải là tạo ra những bước nhảy khổng lồ, mà là các khám phá hay những tiến bộ nhỏ.
- ▶ Do yêu cầu phải tạo ra tính mới, nhà nghiên cứu phải chấp nhận rủi ro, bao gồm:
  - ▶ Không chắc chắn về vấn đề đặt ra là đúng hay sai
  - ▶ Phương pháp nghiên cứu có phù hợp không
  - ▶ Rất nhiều trường hợp phải dừng do không có cách giải quyết phù hợp (dữ liệu, mô hình, bối cảnh)

## Bạn đã sẵn sàng chưa?

- ▶ Nghiên cứu định lượng thách thức hơn, nhưng cũng bổ ích và thú vị hơn.
- ▶ Thường xuyên cảm thấy có quá nhiều khó khăn, rắc rối, và không chắc chắn nếu mô hình hay kết quả có đáng để mất thời gian nghiên cứu.
- ▶ Tất cả các bài nghiên cứu (đặc biệt là chạy mô hình) đều mất nhiều thời gian và khó hơn nhiều so với dự tính ban đầu.
- ▶ Chấp nhận thay đổi quan điểm khi kết quả không đúng với định kiến có trước.

# Cấu trúc phổ biến của một bài nghiên cứu định lượng

- ▶ Giới thiệu đề tài.
  - A Giới thiệu đề tài và câu hỏi nghiên cứu.
  - B Lý do (động cơ) tại sao lại lựa chọn đề tài (hoặc tại sao bạn thấy đề tài này thú vị).
  - C Giới thiệu về phương pháp sử dụng và các kết quả chính.
  - D Giới thiệu về cấu trúc của bài viết/luận văn.
- ▶ Các nghiên cứu đã được thực hiện (Literature review)

Đề cập và giải thích những bài nghiên cứu đã làm có liên quan đến đề tài này.
- ▶ Khung phân tích lý thuyết

Trình bày mô hình hoặc lý thuyết của bài phân tích. Đôi khi phần này được kết hợp với phần Literature review.

- ▶ Ước lượng mô hình và kết quả.
  - A Dữ liệu và thông kê mô tả.
  - B Mô hình và kết quả ước lượng.
  - C Diễn giải kết quả, kiểm định các giả thuyết.
  - D Kiểm định độ vững và độ nhạy của kết quả dựa trên những giả định khác nhau.
  
- ▶ Kết luận
  - A Tái khẳng lại đề tài và câu hỏi nghiên cứu.
  - B Đưa ra câu trả lời cho đề tài nghiên cứu, và có thể so sánh ngắn gọn với những kết quả đã thực hiện.
  - C Chỉ ra điểm mạnh điểm yếu, hướng phát triển sắp tới.
  
- ▶ Tài liệu tham khảo và mục lục

## Giới thiệu đề tài - Introduction

- ▶ Cung cấp cái nhìn khái quát về đề tài, động cơ nghiên cứu, và phương pháp sử dụng, dữ liệu nghiên cứu, và đóng góp của đề tài vào học thuật.
- ▶ Phải đơn giản, dễ hiểu ngay cả khi đọc lần đầu tiên, kể cả đối với những người không thuộc lĩnh vực chuyên môn của đề tài. Tuyệt đối tránh viết quá kỹ thuật, chỉ phù hợp cho một vài chuyên gia. Hầu hết người đọc chỉ đánh giá một bài viết thông qua phần tóm tắt, giới thiệu, và kết luận.
- ▶ Một số nghiên cứu kết hợp phần giới thiệu với phần các nghiên cứu đã thực hiện (literature review), hoặc tách rời.
- ▶ Introduction và literature review là phần *rất khó viết* trong một bài nghiên cứu!

Ví dụ về câu hỏi nghiên cứu và xây dựng giả thuyết cho luận văn tựa đề "Tác động kinh tế của biến đổi khí hậu đến thu nhập trồng lúa ở Việt Nam" của học viên Phạm Thị Kim Phụng (2012).

- ▶ Tác động của khí hậu thông thường hiện nay đến thu nhập của hộ trồng lúa là như thế nào?
- ▶ Dự báo mức độ thiệt hại đối với hộ trồng lúa bị ảnh hưởng bởi BĐKH dựa theo kịch bản BĐKH là bao nhiêu?
- ▶ Giải pháp nào nhằm giảm thiểu mức độ thiệt hại đối với hộ trồng lúa?

Giả thuyết là lời đáp cho những câu hỏi trên, có thể kiểm định được. Ví dụ:

- ▶ Giả thuyết 1: Năng suất lúa gạo phụ thuộc nhiều nhất vào thay đổi nhiệt độ trong thời điểm đơm bông.
- ▶ Giả thuyết 2: Năng suất lúa gạo tăng vào nửa đầu thế kỷ 21 khi nhiệt độ trung bình tăng dần đến ngưỡng nhiệt độ tối ưu. Tuy nhiên nửa sau thế kỷ 21 tác động của nhiệt độ sẽ làm giảm năng suất lúa gạo.



## Các nghiên cứu đã thực hiện - Literature review (LitRev)

- ▶ Mục đích là đặt bài nghiên cứu của bạn trong bức tranh tổng thể của những vấn đề có liên quan. Cần thiết phải làm rõ mối liên hệ giữa một nghiên cứu sâu trong một lĩnh vực cụ thể và vấn đề nghiên cứu lớn (ví dụ nghiên cứu ảnh hưởng của nhiệt độ tăng lên năng suất lúa gạo với ảnh hưởng của biến đổi khí hậu nói chung.)
- ▶ Giải thích các nghiên cứu đã thực hiện để cho thấy tại sao đề tài của bạn là một đóng góp mới cho tổng thể tri thức. Để làm được điều này bạn không chỉ tóm tắt các nghiên cứu đã làm, mà còn phải chỉ ra thiếu sót (trong một số trường hợp). Tuy nhiên, kể cả khi bạn tự tin về kết quả của nghiên cứu của mình, tránh sử dụng từ ngữ quá tiêu cực khi nói về những nghiên cứu của người khác.

- ▶ Tránh viết LitRev theo kiểu chỉ số ("card index"), ví dụ A nói thể này, B nói thể kia, C lại cho rằng...
- ▶ Đôi khi LitRev được viết rất chi tiết. Có thể lập bảng tổng hợp (xem tài liệu tham khảo).
- ▶ Viết LitRev không chỉ là copy và paste lại từ ngữ của người khác mà phải diễn giải lại theo hiểu biết của bạn. Phải hết sức thận trọng khi trích dẫn.
  - ▶ Đọc không hiểu → Trích dẫn nguyên văn (" ") → Biểu hiện của sự lười biếng hoặc thiếu nghiêm túc.
  - ▶ Viết lại ý người khác nhưng trích dẫn sai quy tắc → Đạo văn!

## Các bước để viết LitRev:

- ▶ Phân biệt những cách tiếp cận khác nhau của cùng một vấn đề (ví dụ nếu nói về đánh giá chính sách, có một loạt các phương pháp như đã đề cập trong môn KTL ứng dụng), những giả định, yêu cầu dữ liệu và phương pháp ước lượng.
- ▶ Một số nghiên cứu cần được giải thích sâu hơn những nghiên cứu khác, tùy bối cảnh và nội dung.
- ▶ Nêu ra những điểm mạnh và điểm yếu của những nghiên cứu liên quan, nhưng không quá tiêu cực, và đặt nghiên cứu của bạn trong bối cảnh đó để thấy đâu là đóng góp của bạn.
- ▶ Có thể đọc báo chí điện tử, blogs để xác định vấn đề nghiên cứu, nhưng những nguồn đó không thể dùng được trong phần LitRev. Cần đọc các tạp chí nghiên cứu chuyên ngành.

# Khung lý thuyết phân tích - Theoretical framework

- ▶ Nghiên cứu định lượng vẫn cần phải có khung phân tích lý thuyết!
  - ▶ Mô hình định lượng đôi khi chỉ xác lập mối quan hệ thống kê giữa hai biến số, nhưng không giúp giải thích về bản chất của ước lượng, hay các vấn đề chính sách gặp phải.
  - ▶ Mô hình lý thuyết góp phần làm rõ về lựa chọn mô hình hồi quy (tuyến tính hay phi tuyến), lựa chọn biến nào và bỏ biến nào, tại sao dùng logarithm trong một số trường hợp.
  - ▶ Đôi khi kết hợp khung lý thuyết với các phần khác (phần mô hình và dữ liệu, hoặc LitRev).

# Mô hình và dữ liệu

- ▶ Nguồn gốc của dữ liệu.
- ▶ Mô tả dữ liệu bằng bảng biểu, đồ thị, và giải thích.
- ▶ Giải thích ưu/nhược điểm của dữ liệu.

Viết mô hình ước lượng:

$$\ln(INCOME_i) = \beta_0 + \beta_1 EDUC_i + \beta_2 EXP_i + \beta_3 EXP_i^2 + \beta_4 SEX_i + \varepsilon_i$$

thay vì

$$Y_i = \beta_0 + \beta_1 X1_i + \beta_2 X2_i + \beta_3 X3_i^2 + \beta_4 X4_i + \varepsilon_i$$

Giải thích rõ các ký tự sử dụng, ví dụ  $\varepsilon_i$  phân phối chuẩn iid  $N(0, \sigma^2)$  có nghĩa là gì.

# Giải thích kết quả

- ▶ Giải thích các tham số ước lượng và ý nghĩa thông kê, ý nghĩa kinh tế.
- ▶ Giải thích các kiểm định. Với các kiểm định không phổ biến, cần nêu rõ trị kiểm định, phân phối, bậc tự do, giả thuyết Ho, Ha là gì.
- ▶ Các quy ước làm tròn số.

# Kết luận

- ▶ Tái khẳng định lại mục đích nghiên cứu, những kết quả chính từ việc khảo sát những nghiên cứu đã thực hiện, dữ liệu và phương pháp/mô hình sử dụng để tạo ra kết quả. Có thể so sánh kết quả mới được tìm ra với những nghiên cứu trước đó.
- ▶ Nêu vắn tắt những hạn chế, và đề xuất sử lý hay mở rộng bài nghiên cứu trong tương lai.

## Một số vấn đề đáng lưu ý

- ▶ Chọn phương pháp và mô hình trước hay chọn vấn đề nghiên cứu trước?
- ▶ Viết tiếng Anh hay tiếng Việt?
- ▶ Luận văn là một tài liệu khoa học, không viết như viết báo. Tuyệt đối không viết sáo ngữ. Lập luận phải chặt chẽ, ngôn ngữ rõ ràng, cấu trúc liên kết. Kết luận và khuyến nghị thống nhất với kết quả phân tích.
- ▶ Cung cấp đầy đủ chi tiết về dữ liệu và mã chương trình để người khác mô phỏng lại kết quả nếu bạn dự định xuất bản nghiên cứu.



# Đánh giá một bài nghiên cứu định lượng

Một bài nghiên cứu bắt buộc phải đạt được các yêu cầu sau:

- ▶ Nhận thức rõ ràng về vấn đề nghiên cứu và mối quan hệ với bức tranh tổng thể và các nghiên cứu đã thực hiện.
- ▶ Động cơ rõ ràng của các câu hỏi nghiên cứu được nêu ra và nền tảng vững chắc của mô hình và lý thuyết được sử dụng.
- ▶ Giải thích rõ ràng về phương pháp/mô hình sử dụng và các kết quả chính. Bài viết cũng có thể đề cập đến khả năng mở rộng nghiên cứu nếu phù hợp.
- ▶ Thể hiện rõ ràng giá trị của bài nghiên cứu và những hạn chế của kết quả nghiên cứu.

# Các tiêu chí đánh giá một số bài nghiên cứu định lượng

Các tiêu chí *nội dung* ban đầu:

- ▶ Vấn đề nghiên cứu có được diễn giải rõ ràng không?
- ▶ Các giả thuyết và các giả định có rõ ràng không?
- ▶ Mối quen hệ với các nghiên cứu đã thực hiện có rõ ràng không?
- ▶ Có nêu rõ những hạn chế của bài viết không?
- ▶ Có mô tả dữ liệu và ghi rõ nguồn gốc của dữ liệu không?
- ▶ Phát hiện có ghi có rõ ràng trong phần kết luận không?

Các tiêu chí *hình thức* ban đầu:

- ▶ Có đầy đủ các chương/mục chưa?
- ▶ Font chữ, cách đánh số chương/mục có thống nhất không?
- ▶ Trích dẫn: kiểm tra có chính xác không, bao gồm cả số trang, tên tác giả, và năm xuất bản. Có sắp xếp đúng thứ tự chưa?
- ▶ Bảng biểu có đầy đủ tên, chú thích, đơn vị tính chưa?
- ▶ Phương trình hồi quy có đầy đủ và rõ ràng, đặc biệt với biến chính sách chưa? Có giải thích ký hiệu đầy đủ chưa?

# Báo cáo phản biện mẫu của một bài nghiên cứu định lượng

## Referee Report Rubric

Author's Name: _____	Paper Title: _____	F	Poor	OK	Good	Super
Title and Abstract						
Introduction						
Clear statement of research question						
Literature Review						
Content of lit review						
Organization of lit review						
Citation usage and style						
Theoretical Analysis (may be in lit review section)						
Discussion of underlying theory						
Empirical Analysis						
The Data						
Variable Description and Summary Data Table						
Description of procedure for data collection						
Data limitations and problems						
Empirical Results						
Discussion of DGP/error term (e.g., CEM or Random Xs)						
Regression results table						
Rounding numbers and ease of reading display						
Interpreting Results						
Explanation of a particular coefficient						
If Dummy Dependent Variable, predicted probabilities						

If Dummy Dependent Variable, predicted probabilities					
Including units in discussion of variables					
Appropriate functional forms used and explained					
Discussion of violations of CEM					
Testing for heteroskedasticity or autocorrelation					
Use of robust SEs (if heteroskedasticity present)					
Hypothesis test on estimated coefficient (if applicable)					
Confidence interval on estimated coefficient (if applicable)					
Discussion of economic importance					
Computation and explanation of elasticity (if applicable)					
Comparison of results to previous work (from lit review)					
Answering the research question					
Conclusion					
Summary of paper					
Future work					
References: Content and Style					
Overall Aspects of the Paper					
Content					
Writing					
Style (figures, tables, look and feel)					

Best Parts of the Paper:

Final Evaluation:

Things That Need Work:

Reviewer's Name: \_\_\_\_\_

# Các bước khởi động một bài nghiên cứu thực nghiệm

- ▶ Xác định lĩnh vực nghiên cứu mong muốn (macro, micro, development, environment, trade)
- ▶ Literature review
  - ▶ Bài báo trên các tạp chí chuyên ngành: ngắn gọn súc tích nhưng khó đọc, có thể cung cấp dữ liệu để mô phỏng lại nghiên cứu.
  - ▶ Sách giáo trình: mô hình chuẩn nhưng không có tính sáng tạo.
- ▶ Yêu cầu tiếp cận dữ liệu ở mức độ nào?
  - ▶ VHLSS, VARHS, SME, Census, DHS, WB, GSO...
  - ▶ Nguồn thứ cấp khác.
  - ▶ Dữ liệu tự điều tra.
- ▶ Chạy hồi quy dự báo (agnostic regression, thiết lập mô hình  $Y = f(X_1, X_2, \dots, X_n)$  bằng OLS chỉ dựa trên cảm tính) để xác định yêu cầu dữ liệu có đạt được không, mô hình có hợp lý không trước khi đi sâu vào thiết lập những mô hình chi tiết hơn phù hợp với mục đích của nghiên cứu.

# Thực hành mô phỏng lại các nghiên cứu đã được thực hiện

Là cách tốt nhất để học các kỹ năng về đọc hiểu, lập trình tính toán mô phỏng lại các kết quả đã được công bố!

- ▶ Các tạp chí nổi tiếng thường cung cấp dữ liệu và program code: AER, QJE, JSTOR, Journal of Applied Econometrics Data Archive (<http://qed.econ.queensu.ca/jae/>) và rất nhiều tạp chí khác, website của một số tác giả nổi tiếng để mô phỏng lại các nghiên cứu đã làm.
- ▶ Kiểm tra tính đúng đắn và tính nhạy cảm của kết quả nếu thay đổi các giả định của mô hình:
  - ▶ Các biến có phù hợp không? có vấn đề lựa chọn mẫu (data selection) hay không?
  - ▶ Giả định về cấu trúc hàm có phù hợp không? thay đổi thì dẫn đến hiện tượng gì?

# Thiết kế câu hỏi nghiên cứu như thế nào?<sup>1</sup>

- ▶ Bắt đầu bằng một câu hỏi lớn, ví dụ:
  - ▶ What explains labor force participation?
  - ▶ Are minority women discriminated against in the labor market?
  - ▶ What is the effect of price, income, and education on smoking behavior?
- ▶ Nhiều khả năng (có thể là chắc chắn) câu hỏi đó đã được trả lời trong một nghiên cứu nào đó.
- ▶ Thu hẹp câu hỏi lớn thành những câu hỏi nhỏ nhưng phù hợp với bối cảnh bạn muốn nghiên cứu:
  - ▶ Áp dụng mô hình cũ cho một bộ dữ liệu mới.
  - ▶ Áp dụng cho những vấn đề bạn phát hiện chưa được xử lý trong nghiên cứu ban đầu.

---

<sup>1</sup>Có tham khảo Writing an Empirical Economics Paper, Wabash College

# Cải thiện một bài nghiên cứu như thế nào?

- ▶ Trình bày tại các hội thảo là cách tốt nhất để cải thiện bài viết:
  - ▶ Trình bày sao cho những người chưa bao giờ đọc bài của bạn, hay những người không có chuyên môn có thể hiểu được.
  - ▶ Tạo các quan hệ cần thiết sau này.
- ▶ Gửi bài đi các tạp chí nghiên cứu → Quá trình đánh giá peer-review.



# Tham khảo một số luận văn MPP gần đây

- ▶ Trương Ngọc Phong (2015). Đánh giá khả năng áp thuế tiêu thụ đặc biệt lên mặt hàng nước giải khát công nghiệp tại Việt Nam.
- ▶ Phạm Thị Phương Thúy (2016). Phân tích hiệu quả hoạt động của các công trình cấp nước sinh hoạt nông thôn trên địa bàn tỉnh Đắk Lắk giai đoạn 2010-2014.
- ▶ Nguyễn Ngọc Duy Tuệ (2015). Tác động của trợ cấp đến hoạt động nghiên cứu và phát triển của doanh nghiệp.
- ▶ Hoàng Thị Chinh Thon (2015). Tác động của ô nhiễm tại các làng xã đến tình trạng khám chữa bệnh ở Việt Nam.
- ▶ Phan Văn Hoàng Sơn (2016). Tác động của đầu tư công nghệ thông tin đến hiệu quả hoạt động của doanh nghiệp vừa và nhỏ.
- ▶ Lê Bá Anh (2016). Tác động của vốn vay đến lao động trong doanh nghiệp sản xuất vừa và nhỏ.
- ▶ Chu Phạm Đăng Quang (2017). Tác động của chính sách tăng thuế bảo vệ môi trường đối với xăng lên tiêu dùng xăng trong hoạt động giao thông tại các đô thị Việt Nam.
- ▶ Đặng Thị Ánh Dương (2017). Tác động của việc thay đổi chính sách thuế tiêu thụ đặc biệt đối với mặt hàng xăng lên sản xuất nông nghiệp tại Việt Nam.
- ▶ Hà Diệu Linh (2017). Giá trị sử dụng nước tưới và chính sách giá thủy lợi ở Việt Nam.

## Tài liệu tiếng Anh:

- ▶ Creedy, John 2002. Starting research and writing a research paper, IFLIP Guide 02-1, International Labor Organization.
- ▶ Writing an Empirical Economics Paper, Wabash College.

## Phần II: Các kỹ thuật trong nghiên cứu định lượng

- ▶ Trình bày và sử dụng số liệu (descriptive statistics, data visualization).
- ▶ Những vấn đề có thể dẫn đến sai sót: hiệu lực nội vi và hiệu lực ngoại vi (internal vs external validity); nhân quả và tương quan (causation vs association).
- ▶ Xử lý vấn đề omitted variables bias, incomplete data, spatial autocorrelation, endogeneity/simultaneity, missing data.
- ▶ Các phương pháp đánh giá tác động chính sách (Randomization, PSM, DiD, IV/RD, Heckman's selection model).
- ▶ Kiểm chứng chéo (cross-validation), mô phỏng (bootstraps).
- ▶ Mô hình với biến định tính/biến có giới hạn (limited dependent variables).

- ▶ Mô hình cấu trúc và mô hình rút gọn, hệ phương trình đồng thời (structural vs reduced form model, simultaneous equation systems - SEM).
- ▶ Mô hình cân bằng tổng quát (Computable General Equilibrium - CGE).
- ▶ Các mô hình đánh giá giá trị phi thị trường trong kinh tế học môi trường (hedonic regression, willingness to pay/accept, travel cost method, choice modeling).
- ▶ Phân tích phúc lợi - welfare analysis (distributional equity, economic efficiency).
- ▶ Các xu hướng mới trong nghiên cứu định lượng: big data, GIS/spatial data/spatial econometrics, agent-based simulation model, experimental and behavioral economics, nonparametric models.