

Đói nghèo và bất bình đẳng ở Việt Nam: Các yếu tố về địa lý và không gian

Nicholas Minot (IFPRI)
Bob Baulch (IDS)
và
Michael Epprecht (IFPRI)

phối hợp với nhóm tác chiến lập bản đồ đói nghèo liên Bộ

Thông tin liên lạc:

Ông Đặng Kim Sơn, Giám đốc Trung tâm Thông tin Nông nghiệp và Phát triển Nông thôn/Bộ Nông nghiệp và Phát triển Nông thôn/ Số 2- Ngọc Hà/Hà Nội/ Điện thoại: 84-4-7333895/ Email: Icard1@hn.vnn.vn

Ông Nicholas Minot/Viện Nghiên cứu Chính sách lương thực Quốc tế /2033 K St/Washington D.C.20006 USA/Điện thoại: 1 202 862-5600/Fax: 1 202 467-4439/Email: n.minot@cgiar.org

Ông Bob Baulch/ Viện Nghiên cứu Phát triển/Đại Học Sussex/Brighton BN1 9RE United Kingdom/ Điện thoại: 441273-678774/ Fax: 441273-621202/ Email: b.baulch@ids.ac.uk

Ông Michael Epprecht/ Viện Nghiên cứu Chính sách lương thực Quốc tế/ Số 2- Ngọc Hà/Hà Nội/ Điện thoại: 844-7336508/ Fax: 844-7336507/ Email: Michael@epprecht.org

Bản quyền © 2003 Viện Nghiên cứu Chính sách lương thực Quốc tế và Viện nghiên cứu Phát triển.
Các phần của báo cáo này có thể tái bản không cần giấy phép phát hành nhưng cần có sự chấp nhận của Viện Nghiên cứu Chính sách lương thực Quốc tế và Viện nghiên cứu Phát triển.

Liên lạc: ifpri-copyright@cgiar.org để có thể xin tái bản

Thiết kế bìa: Michael Epprecht.

Lời nói đầu

Trong 15 năm qua, quá trình đổi mới đã đạt được những thành tựu đáng kể trong việc tăng thu nhập và giảm đói nghèo ở Việt Nam. Tuy nhiên, số người nghèo vẫn còn rất nhiều và giúp họ thoát nghèo là một trong những ưu tiên hàng đầu của chính phủ.

Để có thể giải quyết vấn đề khó khăn này thì cần phải trả lời hai câu hỏi: 1) đâu là nguyên nhân của đói nghèo? và 2) những chính sách và đầu tư công cộng nào có thể tấn công các nguyên nhân gây ra đói nghèo? Có rất nhiều cách để trả lời câu hỏi này, một trong số đó là thực hiện phân tích chính sách bằng cách sử dụng bản đồ nghèo đói. Bản đồ nghèo đói kết hợp các số liệu của điều tra hộ gia đình và Tổng điều tra dân số để xây dựng các bản đồ nghèo đói ở các cấp khác nhau. Phương pháp này phát triển trong 5 năm qua và được áp dụng ở trên 12 nước bởi các nhà nghiên cứu của Ngân hàng Thế giới, Viện Nghiên cứu Chính sách Lương thực Quốc tế, và các tổ chức quốc tế khác.

Trước dự án này, phương pháp lập bản đồ nghèo đói được áp dụng 2 lần ở Việt Nam và cho những kết quả khá tốt. Dự án này, “Bản đồ nghèo đói và tiếp cận thị trường ở Việt Nam,” dựa trên các nghiên cứu trước nhưng có những phân tích mới và những ước lượng về tỷ lệ đói nghèo ở cấp thấp hơn. Mục tiêu của dự án là:

- Xây dựng các bản đồ đói nghèo cấp tỉnh, huyện và xã;
- Đánh giá tác động của các yếu tố về nông nghiệp, khí hậu và tiếp cận thị trường tới đói nghèo;
- Nâng cao năng lực cho các tổ chức Việt Nam nhằm xây dựng các bản đồ nghèo đói và phân tích GIS sau này ; và
- Tăng cường sự phối hợp giữa các Bộ trong các vấn đề có liên quan tới nhiều Bộ chức năng như vấn đề phân tích đói nghèo.

Dự án này được tài trợ bởi Cơ quan Phát triển Quốc tế New Zealand (NZAID) với sự hỗ trợ của Ngân hàng Thế giới và Cơ quan Hợp tác Phát triển Thụy Sĩ (SDC). Trong quá trình thực hiện dự án, có sự phối hợp chặt chẽ và hiệu quả của các chuyên gia của Viện Nghiên cứu Chính sách Lương thực Quốc tế (IFPRI) và Viện Nghiên cứu Phát triển (IDS), và các nhà nghiên cứu từ Bộ Kế hoạch và Đầu tư, Bộ Nông nghiệp và Phát triển Nông thôn, Bộ Lao động Thương binh và Xã Hội, Bộ Tài chính, Tổng Cục Thống kê. Sự tham gia của bên Việt Nam còn có Ban điều hành, đứng đầu là Bộ Nông nghiệp và Phát triển Nông thôn, hướng dẫn việc thực hiện dự án, thành lập Nhóm tác chiến lập bản đồ đói nghèo liên Bộ tham gia phân tích và xây dựng bản đồ.

Dự án bắt đầu hoạt động vào giữa năm 2002 với giai đoạn nâng cao năng lực qua 3 đợt tập huấn mỗi đợt kéo dài một tuần cho khoảng 25 nhà phân tích từ bốn Bộ, Tổng Cục Thống kê và các cơ quan nghiên cứu khác. Giai đoạn phân tích triển khai vào đầu năm 2003 với việc thành lập Nhóm tác chiến lập bản đồ đói nghèo liên Bộ và sự tham gia của họ vào các phân tích và xây dựng bản đồ. Các kết quả của dự án được trình bày tại một vài hội thảo, hội thảo cuối cùng vào ngày 2/10/2003 tại Hà Nội.

Báo cáo này sẽ giới thiệu về thông tin nền, mô tả phương pháp nghiên cứu và trình bày các kết quả của dự án. Báo cáo sẽ cung cấp cho người đọc những bức tranh chung về phân bố nghèo đói và các biến liên quan đến đói nghèo ở Việt Nam. Đây là tài liệu tham khảo khá tốt cho các nhà phân tích chính sách, nhà nghiên cứu, giảng viên, các nhà tài trợ và những người làm việc trong lĩnh vực có liên quan đến đói nghèo ở Việt nam.

Tiến sỹ. Đặng Kim Sơn,

Giám đốc Trung tâm Thông tin Nông nghiệp và Phát triển Nông thôn

Bộ Nông nghiệp và Phát triển Nông thôn

Số 2 Ngọc Hà

Hà Nội

Ngày 9 Tháng 12 năm 2003

Lời cảm ơn

Báo cáo được hoàn thành với sự giúp đỡ của rất nhiều các chuyên gia Việt Nam và quốc tế. Ngân sách của dự án được tài trợ bởi Cơ quan Phát triển Quốc tế Thụy sỹ (NZAID). Chúng tôi muốn gửi lời cảm ơn tới đại diện của New Zealand vì những giúp đỡ quý báu của họ:

- Ông Claasen, John (Quản lý Chương trình Cắm Phu Chia/Lào/Thái Lan/Việt Nam, Cơ quan Phát triển Quốc tế Thụy sỹ, Wellington)
- Ông Richard Graves (Phó Đại diện, Đại sứ quán New Zealand, Hà Nội)
- Ông Martin Welsh (Phó Đại diện, Đại sứ quán New Zealand, Hà Nội)

Bên cạnh đó, Ngân hàng Thế giới hỗ trợ dự án qua các hình thức khác nhau, tài trợ nghiên cứu về đói nghèo trước đó dựa trên mẫu 3% của Tổng điều tra Dân số năm 1999, mua số liệu của Tổng điều tra sử dụng cho dự án, và giúp đỡ trong việc thiết kế và điều phối dự án. Đặc biệt, chúng tôi xin gửi lời cảm ơn tới hai người. Bà Nisha Agrawal giúp đỡ trong việc xây dựng ý tưởng phân tích bản đồ đói nghèo. Và ông Mr. Rob Swinkels giúp đỡ chúng tôi trong việc thiết kế dự án, cố vấn cho dự án và tổ chức hội thảo phổ biến các kết quả vào tháng 10/2003. Cơ quan Thụy sỹ về Hợp tác và Phát triển (SDC) đã cử chuyên gia là ông Michael Epprecht tham gia dự án từ ban đầu và là một trong tác giả chính của dự án.

Dự án được thực hiện bởi Viện Nghiên cứu Chính sách Lương thực Quốc tế (IFPRI) và Viện Nghiên cứu Phát triển (IDS) với sự phối hợp của Nhóm tác chiến lập bản đồ đói nghèo liên Bộ và sự chỉ đạo của Ban điều hành.

Ban điều hành gồm các đại diện của các cơ quan hợp tác chính: Bộ Kế hoạch và Đầu tư, Bộ Nông nghiệp và Phát triển Nông thôn, Bộ Lao động Thương binh và Xã hội, Bộ Tài chính, Tổng Cục Thống kê. Đặc biệt, các cá nhân sau đóng góp nhiều ý và chỉ đạo thực hiện dự án:

- Ông Đặng Kim Sơn, Giám đốc, Trung tâm Thông tin Nông nghiệp và PTNT (ICARD), Bộ Nông nghiệp và PTNT (MARD)
- Ông Nguyễn Hải Hữu, Vụ trưởng, Vụ Bảo trợ Xã hội, Bộ Lao động thương binh và xã hội (MOLISA)
- Ông Phạm Hải, Vụ Trưởng, Vụ Địa Phương, Bộ Kế hoạch và Đầu tư (MPI)
- Ông Trương Hùng Long, Vụ Trưởng vụ Ngân sách, Bộ Tài Chính (MOF)
- Ông Nguyễn Phong, Vụ Trưởng, Vụ Môi trường và Xã hội, Tổng cục Thống kê (GSO)

Chúng tôi xin chân thành cảm ơn Tiến sỹ Đặng Kim Sơn, giám đốc Trung tâm Thông tin Nông nghiệp và PTNT (ICARD), trưởng Ban điều hành vì sự hỗ trợ của ông đối với dự án, sự giúp đỡ của ông trong việc tổ chức các lớp tập huấn và cuộc hội thảo cuối cùng, sự chỉ đạo của ông trong Ban điều hành. Sự giúp đỡ của ông thành nhân tố chính đối với sự thành công của dự án.

Nhóm tác chiến liên Bộ được thành lập để phối hợp với các chuyên gia quốc tế trong giai đoạn phân tích của dự án. Đặc biệt, họ phân tích ra các kết quả về tỷ lệ đói nghèo cấp tỉnh, huyện và xã, và xây dựng 7 bản đồ đói nghèo của 7 vùng kinh tế, 61 bản đồ đói nghèo cấp tỉnh trên đĩa CD-ROM. Nhóm tác chiến gồm 12 nhà phân tích từ MARD, MOLISA, MPI, MOF, GSO, Đại học Kinh tế Quốc dân, Viện Kinh tế Nông nghiệp. Thành viên của Nhóm tác chiến lập bản đồ đói nghèo gồm:

- Ông Nguyễn Việt Cường, Đại học Kinh tế Quốc dân
- Ông Đỗ Anh Kiêm, GSO
- Ông Trần Công Thắng, ICARD, MARD
- Ông Trương Hữu Chí, ICARD, MARD
- Ông Nguyễn Ngọc Quế, ICARD, MARD
- Bà Lò Thị Đức, GSO
- Bà Phạm Minh Thu, ILSSA/MOLISA
- Ông Đào Quang Vinh, MOLISA
- Bà Lê Minh Giang, MOLISA
- Ông Nguyễn Chiến Thắng, Viện Kinh tế học
- Ông Nguyễn Việt Hùng, Bộ Tài chính
- Bà Nguyễn Thị Thanh Nga, MPI

Bà Lê Thị Phi Vân, điều phối văn phòng IFPRI tại Hà Nội (công tác tại Viện Kinh tế Nông nghiệp) hỗ trợ rất nhiều trong công tác hậu cần, kế toán, quản lý văn phòng, trợ lý nghiên cứu và liên lạc với Nhóm tác chiến liên Bộ. Bà Lê Thị Phi Vân và ông Trần Công Thắng (ICARD, MARD) dịch báo cáo này sang tiếng Việt từ bản tiếng Anh.

Cuối cùng, chúng tôi xin cảm ơn những người đã đóng góp ý kiến tại hai cuộc hội thảo trình bày kết quả, một ở Trung tâm Hội nghị Quốc tế Hà Nội ngày 2/10/2003 và một ở Trung tâm Thông tin Phát triển Việt Nam (VDIC) cũng tại Hà Nội ngày 3/10/2003. Những người này

gồm có tiến sỹ Đặng Kim Sơn (Giám đốc, ICARD, MARD), tiến sỹ Nguyễn Hải Hữu (Vụ trưởng, Vụ Bảo trợ Xã hội, MOLISA), ông Nguyễn Phong (Vụ trưởng, Vụ Xã hội Môi trường, GSO), và ông Rob Swinkels (Ngân hàng Thế giới, Hà Nội).

Chúng tôi hy vọng các kết quả của nghiên cứu này sẽ giúp ích phần nào cho Chính phủ Việt Nam, các tổ chức phi chính phủ, và các tổ chức quốc tế trong việc xây dựng các chính sách và các chương trình xoá đói giảm nghèo ở Việt Nam.

Nicholas Minot, Phòng nghiên cứu

Viện Nghiên cứu Chính sách Lương thực Quốc tế

Bob Baulch,

Viện Nghiên cứu Phát triển

Michael Epprecht,

Viện Nghiên cứu Chính sách Lương thực Quốc tế

12 Tháng 12 năm 2003

Tóm tắt

Báo cáo này sử dụng phương pháp tương đối mới gọi là “Phương pháp ước lượng diện tích nhỏ” để ước lượng các chỉ số đói nghèo và bất bình đẳng ở cấp xã, huyện và tỉnh ở Việt Nam. Phương pháp kết hợp thông tin trong Điều tra mức sống dân Việt nam 1997-1998 và Tổng Điều tra dân số và nhà ở năm 1999

Các kết quả cho thấy tỷ lệ đói nghèo (P0) là cao nhất ở những vùng sâu vùng xa của Tây Bắc, Đông Bắc và phía bắc của Tây Nguyên. Tỷ lệ đói nghèo ở Đồng bằng sông Hồng và Đồng bằng sông Cửu Long là trung bình. Tỷ lệ đói nghèo thấp nhất là ở các thành phố lớn, Hà Nội và thành phố Hồ Chí Minh và vùng Đông Nam Bộ. Độ chính xác của các ước lượng là hợp lý cho cấp huyện và tỉnh, nhưng các kết quả cấp xã cần phải sử dụng cẩn trọng vì một số ước lượng không có độ chính xác cao

Bản đồ về mật độ đói nghèo cho thấy, mặc dù tỷ lệ đói nghèo ở các vùng núi xa xôi là cao nhất nhưng những vùng này dân cư thưa thớt nên hầu hết những người nghèo sống ở Đồng bằng sông Hồng và Đồng bằng sông Cửu Long

So sánh những kết quả này với những ước lượng đói nghèo cấp huyện của MOLISA, chúng tôi thấy có sự khác nhau. Trong báo cáo, chúng tôi đã chỉ ra một số lý do cho sự khác biệt này.

Phân tích này củng cố các nghiên cứu cho rằng bất bình đẳng trong chi tiêu bình quân đầu người là tương đối thấp ở Việt Nam. Bất bình đẳng cao nhất ở những thành phố lớn và (rất ngạc nhiên) ở một số vùng miền núi. Bất bình đẳng thấp nhất ở Đồng bằng sông Hồng, tiếp theo là Đồng bằng sông Cửu Long. Khoảng 2/3 bất bình đẳng do sự chênh lệch trong các huyện chứ không phải giữa các huyện

Tỷ lệ đói nghèo cấp huyện có tương quan chặt chẽ với chi tiêu bình quân đầu người trung bình của huyện. Nói cách khác, bất bình đẳng thay đổi rất ít từ huyện này sang huyện khác. Bất bình đẳng cao hơn ở huyện có chi tiêu bình quân đầu người cao, mặc dù tương quan này không thật sự mạnh. Hơn nữa, bất bình đẳng cao ở khu vực thành thị.

Các biến nông nghiệp và khí hậu và tiếp cận thị trường có thể giải thích 3/4 sự biến động tỷ lệ đói nghèo cấp huyện ở nông thôn. Tỷ lệ đói nghèo cao ở những huyện có độ dốc cao, nhiều núi đá và đất đai kém (cát, mặn và axit sunfuric), và cách xa thành phố. Ngược lại, các biến về khí hậu nông nghiệp không giải thích tốt biến động đói nghèo ở thành thị

Mô hình hồi quy vùng, trong đó các hệ số biến đổi từ vùng này tới vùng khác, cho thấy những vùng đất bằng phẳng và mật độ đường xá cao thì tỷ lệ đói nghèo thấp hơn. Tuy nhiên, các biến khác như lượng mưa, diện tích rừng có thể có tương quan cùng chiều với đói nghèo ở những vùng này nhưng có thể có tương quan ngược chiều ở những vùng khác. Nói chung tương quan giữa các biến về khí hậu nông nghiệp và đói nghèo biến đổi khá lớn giữa các vùng ở Việt Nam.

Nhiều chương trình xóa đói giảm nghèo đều hướng vào những vùng miền núi Việt Nam. Các kết quả trong nghiên cứu này chỉ ra rằng, chúng ta có thể phát triển định hướng của những chương trình này bằng cách áp dụng các ước lượng đói nghèo cấp huyện và xã chính xác hơn, dù cho những nghiên cứu tới cần phải phân tích rõ hơn sự khác biệt giữa những kết quả tính theo các phương pháp khác nhau

Các biến khí hậu nông nghiệp có thể giải thích phần lớn sự khác nhau giữa đói nghèo nông thôn, điều này cho thấy đói nghèo ở vùng sâu vùng xa có sự liên quan chặt chẽ với tiềm năng nông nghiệp hạn chế và thiếu sự tiếp cận thị trường. Phân tích cho thấy tầm quan trọng của nâng cao khả năng phát triển thị trường. Vấn đề thực tế là đói nghèo ở vùng sâu vùng xa có sự liên quan với tiềm năng nông nghiệp hạn chế cho thấy những cố gắng hạn chế di cư khỏi những vùng có điều kiện tự nhiên không thuận lợi có thể không phải là chính sách tốt để xóa đói giảm nghèo.

Cuối cùng, nghiên cứu nêu lên rằng phương pháp ước lượng diện tích nhỏ không hữu ích cho việc xây dựng các bản đồ hàng năm vì phải dựa trên số liệu Tổng điều tra, nhưng phương pháp này có thể sử dụng để biểu thị theo không gian các yếu tố khác mà các nhà làm chính sách quan tâm như đa dạng hóa thu nhập, thặng dư nông nghiệp và khả năng tổn thương. Bên cạnh đó, nó có thể sử dụng để ước lượng tỷ lệ đói nghèo trong nhóm người dễ bị tổn thương tới nhóm nhỏ hơn với số liệu điều tra hộ gia đình như nhóm dân tộc thiểu số nhỏ, hay những người đánh cá.

MỤC LỤC

1. Giới thiệu chung	1
2. Số liệu và phương pháp	5
2.1 Số liệu	5
2.2 Phương pháp ước lượng tỷ lệ đói nghèo	5
2.3 Phương pháp ước lượng các chỉ số đói nghèo khác	10
3. Đói nghèo và bất bình đẳng xét về mặt không gian	19
3.1 Đặc điểm của hộ liên quan đến chỉ tiêu bình quân đầu người	19
3.2 Tỷ lệ đói nghèo	25
3.3. Các chỉ số đói nghèo về không gian khác	41
3.4. Bất bình đẳng giữa các vùng	43
3.5. Mối liên hệ giữa thu nhập, đói nghèo và bất bình đẳng	48
3.6 So sánh với ước lượng về đói nghèo của MOLISA	52
4. Các yếu tố địa lý ảnh hưởng đến đói nghèo	55
4.1. Các yếu tố địa lý	55
4.2. Các vấn đề trong ước lượng	57
4.3. Mô hình tổng thể phân tích đói nghèo nông thôn	58
4.4. Mô hình tổng thể đói nghèo thành thị (global model of rural poverty)	62
5. Biến động về không gian của các yếu tố ảnh hưởng tới đói nghèo	67
5.1. Mô tả mô hình	67
5.2. Các kết quả	69
6. Tóm tắt và kết luận	75
6.1. Tóm tắt	75
6.2. Kết luận	79
6.3 Khuyến nghị về mặt chính sách và cho các chương trình	80
6.4. Định hướng cho nghiên cứu trong tương lai	83
Tài liệu tham khảo	91
Phụ lục A: Sử dụng các biến GIS trong phân tích thống kê	89
Phụ lục B. Định nghĩa thuật ngữ	97

Danh sách bảng

Bảng 1. Đặc điểm hộ gia đình trong Tổng điều tra dân số và trong VLSS	7
Bảng 2. Các biến giả thích sử dụng trong phân tích hồi quy về không gian	14
Bảng 3. Mô tả các biến sử dụng trong mô hình	20
Bảng 4. Mô hình hồi quy log của chi tiêu bình quân đầu người.....	22
Bảng 5. Ý nghĩa thống kê của các nhóm biến.....	24
Bảng 6. Tỷ lệ đói nghèo (P0) thành thị và nông thôn theo tỉnh.....	26
Bảng 7. Phân tích tỷ lệ bất bình đẳng thành các nhóm giữa tỉnh và trong cùng tỉnh	47
Bảng 8. Phân tách sự bất bình đẳng thành các nhóm giữa huyện và trong huyện.....	48
Bảng 9. Các yếu tố khí hậu nông nghiệp và kinh tế xã hội có thể ảnh hưởng tới tỉ lệ đói nghèo.....	55
Bảng 10. Các biến được sử dụng để phân tích các yếu tố địa lý	56
Bảng 11. Kiểm tra sự phụ thuộc không gian trong mô hình đói nghèo cho nông thôn	58
Bảng 12. Mô hình tổng thể các yếu tố địa lý ảnh hưởng tới đói nghèo ở nông thôn.....	59
Bảng 13. Mô hình chọn lọc các yếu tố địa lý ảnh hưởng đến đói nghèo.....	61
Bảng 14. Kiểm định sự phụ thuộc về không gian đối với mô hình đói nghèo thành thị	62
Bảng 15. Mô hình đầy đủ các yếu tố địa lý ảnh hưởng đến đói nghèo thành thị.....	63
Bảng 16. Mô hình lựa chọn phân tích các yếu tố địa lý ảnh hưởng tới đói nghèo thành thị ..	65
Bảng 17. Các biến sử dụng trong phân tích hồi quy quyền sở địa lý.....	67
Bảng 18. Tóm tắt kết quả của mô hình đói nghèo tổng thể	69
Bảng 19. Tóm tắt kết quả của mô hình thể giới và địa phương.....	70
Bảng 20. Kết quả ước lượng tham số từ mô hình địa phương.....	71
Bảng 21. Mức ý nghĩa thống kê của biến đổi không gian trong tham số ước lượng	73

Danh sách các hình

Hình 1. Bản đồ tỷ lệ đói nghèo (P_0) theo tỉnh.....	29
Hình 2. Đồ thị tỷ lệ đói nghèo theo tỉnh và khoảng tin cậy	31
Hình 3. Bản đồ đói nghèo cấp huyện (P_0).....	33
Hình 4. Tỷ lệ đói nghèo huyện và khoảng tin cậy	34
Hình 5. Tỷ lệ đói nghèo của thành thị và nông thôn cấp huyện	35
Hình 6. Bản đồ đói nghèo cấp xã.....	37
Hình 7. Bản đồ về độ cao và đường xá.....	38
Hình 8. Tỷ lệ đói nghèo cấp xã và khoảng tin cậy.....	39
Hình 9. Bản đồ mật độ đói nghèo	40
Hình 10. Bản đồ độ sâu của đói nghèo (P_1) và độ trầm trọng của đói nghèo (P_2) của mỗi huyện	42
Hình 11. Độ sâu của đói nghèo (P_1) và độ trầm trọng của đói nghèo (P_2) iểu diễn như một hàm tỉ lệ đói nghèo (P_0) ở mỗi huyện	42
Hình 12. Bản đồ bất bình đẳng tính theo hệ số Gini.....	44
Hình 13. Bản đồ bất bình đẳng qua chỉ số Theil L và Theil T.....	46
Hình 14. Chỉ số bất bình đẳng Theil L và Theil T là hàm của hệ số Gini cho mỗi huyện	46
Hình 18. Tỉ lệ đói nghèo (P_0) như một hàm số của chỉ tiêu bình quân đầu người.....	49
Hình 19. Hệ số Gini bất bình đẳng như một hàm số của chỉ tiêu bình quân đầu người	50
Hình 20. Hệ số Gini bất bình đẳng như một hàm số của tỉ lệ đói nghèo (P_0).....	50
Hình 20. Hệ số Gini bất bình đẳng như một hàm số của tỉ lệ đói nghèo (P_0).....	50
Hình 21. Tỷ lệ đói nghèo (P_0) là hàm của mức độ đô thị hoá	51
Hình 22. Hệ số bất bình đẳng Gini là hàm của đô thị hoá	52
Hình 23. So sánh tỷ lệ đói nghèo (P_0) từ MOLISA và từ phương pháp ước lượng diện tích nhỏ.....	54
Hình 24. Bản đồ phân bố không gian của các biến độc lập.....	68
Hình 25. Tóm tắt kết quả của mô hình địa phương và thế giới	70
Hình 26. Bản đồ phân bố không gian của R^2 từng vùng.....	71
Hình 27. Bản đồ phân bố không gian các hệ số vùng của các biến độc lập	72

1. Giới thiệu chung

Ở hầu hết các nước, đói nghèo thường tập trung theo không gian. Đói nghèo cao thường ở những vùng xa xôi với các điều kiện thời tiết, đất đai không thuận lợi trong khi ở những vùng gần với thành phố và những thị trường lớn thì có tỷ lệ đói nghèo thấp hơn. Những thông tin về sự phân bố đói nghèo là sự quan tâm của các nhà chính sách và nghiên cứu bởi vì các nguyên nhân sau. Thứ nhất, nó có thể sử dụng để định lượng sự bất cân bằng về mức sống giữa các vùng và xác định những vùng nào bị tụt hậu trong tiến trình phát triển kinh tế. Thứ hai, nó tạo điều kiện thuận lợi đặt ra mục tiêu của các chương trình, ít nhất là các chương trình nhằm xoá đói giảm nghèo như giáo dục, sức khỏe, tín dụng, hỗ trợ lương thực. Thứ ba, nó có thể chỉ ra các nhân tố về địa lý gắn với đói nghèo như địa hình vùng núi hay khoảng cách xa so với các thành phố.

Ở rất nhiều nước, nguồn thông tin chính của đói nghèo xét theo không gian là các cuộc điều tra chi tiêu và thu nhập. Các cuộc điều tra này thường gồm từ 2000 đến 8000 hộ, chỉ cho phép ước tính đói nghèo của 3 đến 12 vùng trong một quốc gia. Các nghiên cứu chỉ ra rằng, việc đặt ra các mục tiêu giảm nghèo theo địa lý là hiệu quả nhất khi các vùng địa lý là khá nhỏ, ví dụ như một xã, một huyện (Baker và Grosh, 1994; Bigman và Fofack, 2000). Các thông tin về hộ ở các cấp này chỉ có thể tìm trong Tổng điều tra dân số, nhưng bảng hỏi của Tổng điều tra thường rất hạn chế về các đặc điểm của hộ và rất ít khi bao gồm các câu hỏi liên quan đến thu nhập và chi tiêu.

Ở Việt Nam, có ít nhất 3 nguồn thông tin liên quan đến tỷ lệ đói nghèo. Thứ nhất, Tổng cục Thống kê (GSO) thực hiện hai cuộc Điều tra mức sống dân cư Việt Nam, một năm 1992-93 và một năm 1997-98. Với số mẫu tương ứng là 4000 hộ và 6000 hộ, các cuộc điều tra này có thể ước lượng tỷ lệ đói nghèo của 7 vùng Việt Nam¹.

GSO cũng thực hiện các cuộc điều tra hộ gia đình lớn hơn, như Điều tra hộ gia đình đa mục tiêu và Điều tra mức sống dân cư Việt Nam năm 2001. Những cuộc điều tra này có số mẫu hàng chục nghìn hộ và có thể ước lượng tỷ lệ đói nghèo chính xác đến cấp tỉnh.

Một nguồn thông tin khác về sự phân bố đói nghèo theo không gian là của Bộ Lao động thương binh và xã hội (MOLISA). Mỗi năm, MOLISA chuẩn bị danh sách của các hộ nghèo ở các xã dựa trên thông tin thu thập của lãnh đạo địa phương sử dụng các tiêu chí xác định nghèo đói của MOLISA. Các chỉ số về phúc lợi là thu nhập bình quân đầu người, nhưng đường nghèo đói lại xác định dựa trên giá trị lượng gạo nhất định theo giá của vùng. Thông tin này được sử dụng để xác định các xã nghèo nhất đủ điều kiện cho vào các chương trình

¹ Trước năm 1998, Việt Nam chia làm 7 vùng: Miền núi và trung du phía bắc, ĐBSH, Bắc trung bộ, Nam trung bộ, Tây nguyên, Đông Nam bộ và ĐBSCL. Năm 1998, hệ thống này được chia lại, Miền núi và trung du phía bắc chia thành Đông Bắc và Tây Bắc, tỉnh Lâm Đồng từ tây Nguyên sang Đông nam Bộ, và tỉnh Ninh Thuận và Bình Thuận từ Nam Trung Bộ sang Đông Nam Bộ.

đặc biệt và được trợ cấp để xóa đói giảm nghèo. Mặc dù hệ thống này không tốn chi phí cao và có thể cung cấp thông tin đói nghèo hàng năm, nhưng các tỉnh lại dựa trên các tiêu chuẩn đói nghèo khác nhau và sử dụng hướng dẫn thu thập thông tin khác nhau cho phân tích. Hơn nữa, thậm chí khi những hướng dẫn này là thống nhất, việc sử dụng hàng nghìn các người thu thập thông tin sẽ rất khó đảm bảo có sự thống nhất trong việc áp dụng các hướng dẫn khi điều tra (xem Conway, 2001).

Những năm gần đây, một công nghệ mới được gọi là "ước lượng diện tích nhỏ" (small area estimation) được phát triển, trong đó kết hợp cả số liệu của tổng điều tra dân số và điều tra hộ gia đình để ước lượng tỷ lệ đói nghèo (hay các chỉ tiêu khác) cho các đơn vị địa lý nhỏ hơn. Mặc dù có rất nhiều cách tiếp cận, những tất cả đều có 3 bước. Bước thứ nhất, lựa chọn các đặc điểm của hộ gia đình ở cuộc điều tra hộ gia đình và tổng điều tra dân số, như thành phần gia đình, giáo dục, nghề nghiệp, đặc điểm nhà ở và sở hữu tài sản. Thứ hai, số liệu ở cuộc điều tra hộ gia đình được sử dụng để ước lượng đói nghèo hay ước lượng chỉ tiêu là một hàm của các đặc điểm hộ gia đình. Thứ ba, số liệu ở tổng điều tra dân số sẽ được cho vào phương trình để ước lượng đói nghèo cho các khu vực địa lý nhỏ hơn.

Ví dụ, Minot (1998 và 2000) sử dụng điều tra mức sống dân cư 1992-93 và mô hình Probit để ước lượng đói nghèo của hộ nông thôn là hàm của một loạt các đặc điểm hộ. Những giá trị trung bình của những đặc điểm này ở cấp huyện cũng được thu thập từ Tổng điều tra nông nghiệp 1994 và được đưa vào trong mô hình nhằm ước lượng tỷ lệ đói nghèo khu vực nông thôn cho 534 huyện.

Elbers và các đồng nghiệp (2003) phát triển phương pháp tương tự trong đó sử dụng số liệu điều tra và tổng điều tra của Ecuador (2000). Bằng cách sử dụng mô hình hồi quy tuyến tính-lo-ga-rít và số liệu hộ gia đình từ tổng điều tra, họ có thể cho thấy rằng phương pháp của họ tạo ra các ước lượng không chệch về tỷ lệ đói nghèo và cũng có thể tính các sai số chuẩn của các ước lượng về đói nghèo². Cách tiếp cận này được áp dụng ở một số nước khác như Panama và Nam Phi (xem Ngân hàng thế giới, 2000; Thống kê Nam Phi và Ngân hàng Thế giới, 2000).

Nghiên cứu về Việt Nam trước đây có một vài hạn chế. Thứ nhất, vì dựa quá nhiều và tổng điều tra Nông nghiệp nên chỉ có thể ước lượng tỷ lệ đói nghèo ở nông thôn. Thứ hai, việc sử dụng hàm probit và các giá trị trung bình cấp huyện, mặc dù trực quan thì có thể tin cậy nhưng không tạo ra các ước lượng không chệch về tỷ lệ đói nghèo cấp huyện (xem Minot và Baulch, 2002b cho ước lượng độ lớn sai số). Thứ ba, do không có số liệu tổng điều tra hộ gia đình nên không thể ước lượng các sai số chuẩn của các chỉ số đói nghèo để đánh giá độ chính xác của các kết quả này.

Gần đây hơn, Minot và Baulch (2002a) sử dụng Điều tra mức sống dân cư Việt nam 1998 và mẫu 3% tổng điều tra dân số để tạo ra các ước lượng về tỷ lệ đói nghèo ở thành thị, nông thôn

của các tỉnh Việt Nam. Không giống các phân tích bản đồ đói nghèo trước, nghiên cứu này sử dụng số liệu Tổng điều tra dân số cấp hộ, cho phép việc tính toán các sai số chuẩn của các ước lượng đói nghèo sử dụng các phương pháp mô tả trong Elbers và các đồng nghiệp (2003) và Hentschel và đồng nghiệp (2000).

Nghiên cứu này có 3 mục tiêu:

- Mô tả đói nghèo và bất bình đẳng giữa các vùng của Việt Nam
- Nghiên cứu các yếu tố địa lý (bao gồm các yếu tố về khí hậu, nông nghiệp và tiếp cận thị trường) tác động tới đói nghèo thành thị và nông thôn Việt Nam
- Phổ biến thông tin về phương pháp ước lượng diện tích nhỏ và ứng dụng trong thời gian tới ở Việt Nam

Báo cáo này phát triển nghiên cứu trước của Minot và Baulch (2002a) theo 2 cách:

- Bằng việc sử dụng mẫu lớn hơn (mẫu 33%) của Tổng điều tra dân số năm 1999, chúng ta có thể ước lượng tỷ lệ đói nghèo ở huyện và xã (các nghiên cứu trước chỉ ước lượng tỷ lệ đói nghèo cấp tỉnh)
- Bên cạnh xác định tỷ lệ đói nghèo, nghiên cứu này cũng tính toán hai chỉ tiêu đói nghèo khác và 3 chỉ tiêu về bất bình đẳng ở cấp huyện.
- Không như các nghiên cứu trước, nghiên cứu này phân tích tác động của các yếu tố địa lý (bao gồm các yếu tố về khí hậu, nông nghiệp và tiếp cận thị trường) tới đói nghèo ở khu vực thành thị và nông thôn

Báo cáo này gồm 6 Chương. Tiếp theo phần giới thiệu chung này, Phần 2 miêu tả phương pháp và số liệu sử dụng trong nghiên cứu này. Phần 3 sẽ xem xét xu hướng không gian của đói nghèo và bất bình đẳng ở Việt Nam sử dụng các chỉ số về đói nghèo và 3 chỉ tiêu về bất bình đẳng. Phần 4 nghiên cứu các yếu tố về địa lý ảnh hưởng tới đói nghèo, sử dụng phương pháp hồi quy không gian và tập hợp các biến lấy từ hệ thống thông tin địa lý (GIS). Phần 5 sẽ nghiên cứu sự biến động về không gian trong tương quan giữa đói nghèo và các yếu tố địa lý sử dụng phép hồi quy theo vùng có quyền số (locally weighted regression). Cuối cùng, Phần 6 sẽ tóm tắt các kết quả và thảo luận một số khuyến nghị về chính sách và nghiên cứu tới.

² tỷ lệ đói nghèo (poverty headcount) được xác định là lệ dân số có chỉ tiêu bình quân đầu người dưới mức nghèo đói.

2. Số liệu và phương pháp nghiên cứu

2.1 Số liệu

Xây dựng bản đồ đói nghèo của dự án này sử dụng hai bộ số liệu: Điều tra mức sống Việt Nam 1998 (VLSS) và Tổng điều tra nhà ở và dân số năm 1999. VLSS thực hiện bởi GSO do cơ quan Phát triển Quốc tế Thụy Sĩ và UNDP tài trợ với sự hỗ trợ về kỹ thuật của Ngân hàng Thế giới. Mẫu điều tra gồm 6000 hộ, tạo thành mẫu phân tầng ngẫu nhiên. Mẫu này gồm 4270 hộ ở nông thôn và 1730 hộ ở thành thị.

Tổng điều tra nhà ở và dân số năm 1999 được thực hiện bởi GSO và đề cập đến thời điểm 1/4/1999. Điều tra này được thực hiện với sự hỗ trợ về kỹ thuật và tài chính của Cơ quan kế hoạch hoá gia đình Liên Hợp quốc và UNDP. Kết quả đầy đủ của Tổng điều tra không phát hành bởi GSO, nhưng chúng tôi có thể tiếp cận mẫu 33% của Tổng điều tra. Mẫu 33% của Tổng điều tra được lựa chọn bởi GSO sử dụng mẫu hệ thống cứ 3 hộ chọn một dựa trên danh sách hộ sắp xếp theo đơn vị hành chính. Mẫu bao gồm 5.553.811 hộ.

Phân phân tích không gian trong dự án này sử dụng các biến liên quan đến không gian địa lý mô tả khí hậu, địa hình, đất đai, địa lý, và tiếp cận thị trường. Dữ liệu về địa hình được lấy từ Trung tâm Dịch vụ Địa chất Liên hợp Quốc (USGS). Dữ liệu về đường xá và biên giới hành chính lấy từ Trung tâm Viễn Thám và Geomatics, thuộc Tổng cục Quản lý đất. Dữ liệu đất đai và khí hậu lấy từ Trung tâm Thông tin Nông nghiệp và Phát triển Nông thôn. Và dữ liệu về dân số và các dữ liệu địa lý khác lấy từ mẫu 33% Tổng điều tra nhà ở và dân số. Nhiều số liệu trong các biến phải làm sạch, kiểm tra, xử lý để có thể tạo ra các biến có thể sử dụng trong phân tích không gian.

2.2 Phương pháp ước lượng tỷ lệ đói nghèo

Đường đói nghèo sử dụng trong nghiên cứu này là "đường đói nghèo chung" sử dụng trong phân tích VLSS 1997-98 và trong báo cáo "Tấn công đói nghèo" (xem GSO, 2000 và Nhóm Chuyên gia về nghèo đói, 1999). Đường nghèo đói tương ứng với mức chi tiêu (bao gồm cả chi tiêu do hộ tự sản xuất và điều chỉnh theo sự khác nhau về giá giữa các vùng và mùa vụ) đủ để cung cấp 2100Kcal/người/ngày sử dụng rõ hàng hoá của hộ ở nhóm thứ 3, cộng với chi tiêu phi lương thực bằng với mức hộ ở nhóm thứ 3 sử dụng chi tiêu cho hàng hoá phi lương thực. Đường nghèo đói là 1.789.871đ/người/năm, nhưng chi tiêu dùng trong điều tra được điều chỉnh bằng cách sử dụng các chỉ số giá vùng theo tháng để bù đắp cho sự khác nhau của các chi phí trong quá trình điều tra và giữa các vùng.

Lập bản đồ đói nghèo là một trong những ứng dụng của phương pháp "ước lượng diện tích nhỏ". Phương pháp này được chia làm ba giai đoạn cụ thể. Giai đoạn 0 liên quan đến việc xác định các biến mô tả các đặc điểm của hộ có thể liên quan tới thu nhập và đói nghèo có trong

cả điều tra hộ gia đình và Tổng điều tra dân số. Giai đoạn 1, đo lường phúc lợi, thường là chỉ tiêu bình quân đầu người, được ước lượng bằng hàm của các đặc điểm hộ sử dụng phân tích hồi quy và số liệu điều tra hộ gia đình. Giai đoạn 2, phương trình hồi quy này được áp dụng cho các đặc điểm tương tự của hộ gia đình trong Tổng điều tra dân số, tính toán phúc lợi (welfare) ước tính cho mỗi hộ trong Tổng điều tra. Thông tin này, sau đó, được tổng hợp tính cho các đơn vị hành chính, như huyện, tỉnh để ước lượng tỷ lệ nghèo đói và sai số chuẩn của các ước lượng.

Ba phần sau đây sẽ miêu tả các phương pháp này chi tiết hơn, cũng như mô tả cách áp dụng cho nghiên cứu này.

Giai đoạn 0: Xác định các đặc điểm của hộ trong VLSS và Tổng điều tra nhà ở và dân số

Bước đầu tiên là so sánh bảng hỏi của VLSS1998 và Tổng điều tra nhà ở và dân số để xác định những đặc điểm nào của hộ có cả trong hai cuộc điều tra và có thể dùng như những chỉ tiêu đói nghèo. Bên cạnh so sánh hai bảng hỏi, cũng cần so sánh giá trị của các biến để đảm bảo rằng những giá trị này miêu tả những đặc điểm giống nhau. Ví dụ, loại việc làm đầu tiên phải quan tâm, những xem xét kỹ hơn cho thấy sự phân chia loại công việc trong VLSS và Tổng điều tra là khác nhau, vì thế số liệu ở hai cuộc điều tra là không trùng khớp. Vì vậy, biến này không thể cho vào phân tích được. Dựa trên sự so sánh này, có 17 đặc điểm của hộ là được lựa chọn trong phân tích lập bản đồ nghèo đói. (xem Bảng 1).

Một số đặc điểm của hộ có tính chất phân loại khi hồi quy phải được biểu thị bởi các biến giả (nhị nguyên). Ví dụ nguồn nước uống chính là một đặc điểm của hộ, nhưng để phân tích hồi quy thì nó phải được trình biểu thị qua các biến giả riêng biệt cho nước máy bên ngoài, vòi nước máy trong nhà, giếng khoan, giếng thường ... Vì vậy 17 đặc điểm của hộ được trình bày trong phân tích hồi quy qua 39 biến.

Giai đoạn 1: Ước tính chỉ tiêu bình quân với điều tra hộ gia đình

Như đề cập ở trên, Giai đoạn 1 của phương pháp lập bản đồ nghèo đói liên quan đến việc sử dụng số liệu điều tra của hộ và phân tích hồi quy để ước lượng phúc lợi của hộ là hàm của các đặc điểm hộ gia đình. Trong nghiên cứu này chúng tôi sử dụng chỉ tiêu bình quân đầu người thực tế trong VLSS1997-98 như chỉ tiêu phúc lợi của hộ. Các biến giải thích là 17 đặc điểm của hộ như miêu tả ở trên, biểu thị qua 39 biến. Học thuyết kinh tế cho thấy không có sự hướng dẫn nào về dạng hàm nhưng nhìn chung hàm logarit - tuyến tính hay được sử dụng:

$$\ln(y_i) = X_i'\beta + \varepsilon_i \quad (1)$$

Bảng 1. Đặc điểm hộ gia đình trong Tổng điều tra dân số và trong VLSS

Những đặc điểm của hộ	Số câu hỏi	
	Tổng điều tra	VLSS
	1999	1998
Số người trong hộ	Pt I,Q4	S1A
tỷ lệ thành viên hộ trên 60 tuổi	Pt I, Q4	S1A,Q2
tỷ lệ thành viên hộ dưới 15	Pt I, Q4	S1A,Q6
tỷ lệ thành viên hộ là phụ nữ	Pt I, Q3	S1A,Q6
Trình độ học vấn của chủ hộ	Pt I,Q11-13	S2A
Chủ hộ có vợ/chồng không	Pt I, Q2	S1B,Q3
trình độ vợ/chồng của chủ hộ	Pt I,Q11-13	S2A
Chủ hộ có thuộc dân tộc thiểu số không	Pt I, Q4	S0A
Nghề nghiệp chính của chủ hộ trong 12 tháng qua	Pt I, Q16	S4D
Loại nhà (kiên cố, bán kiên cố, khung gỗ, "đơn giản")	Pt III, Q3	S6A,Q1
Loại nhà tương quan với khu vực sống	Pt III, 4	S6C,Q1a
Nhà có dùng điện hay không	Pt III, Q7	S6B,Q33
Nguồn nước uống chính	Pt III, 8	S6B,Q25
Loại nhà vệ sinh	Pt III, Q9	S6B,Q31
Nhà có TV không	Pt III, Q10	S12C
Nhà có đài không	Pt III, Q11	S12C
Vùng nơi hộ sống	trang 1	S0A

Nguồn: bảng hỏi VLSS 1998 và Tổng điều tra dân số 1999

trong đó y_i là chỉ tiêu bình quân đầu người thực tế của hộ i , X_i' là vector của các đặc điểm hộ i , β là vector của các hệ số ước lượng, ε_i là số dư ngẫu nhiên phân bố $N(0, \sigma)$. Vì quan tâm của chúng tôi là ước lượng giá trị $\ln(y)$ hơn là đánh giá tác động của mỗi biến giải thích, nên chúng tôi không quan tâm đến sự nội sinh có thể của các biến giải thích. Elbers và đồng nghiệp (2003) chỉ rằng xác suất mà hộ i với đặc điểm X_i là nghèo có thể diễn đạt bằng:

$$E[P_i | X_i, \beta, \sigma^2] = \Phi \left[\frac{\ln z - X_i' \beta}{\sigma} \right] \quad (2)$$

trong đó P_i là biến nhận giá trị 1 nếu hộ là nghèo và 0 nếu ngược lại, z là "đường đói nghèo chung" (xem GSO, 2000, trang 260), và Φ là hàm chuẩn hoá tích lũy cumulative standard

normal function). Nếu giá trị ước lượng (predicted) log chỉ tiêu bình quân đầu người ($X_i'\beta$) bằng với log đường chuẩn nghèo ($\ln(z)$), thì giá trị trong ngoặc sẽ bằng 0 và xác suất mà hộ i là nghèo là 50%. Giá trị chỉ tiêu ước lượng thấp hơn chuẩn nghèo thì giá trị trong ngoặc sẽ lớn hơn 0 và có xác suất hộ là nghèo sẽ lớn hơn, trong khi đó giá trị chỉ tiêu ước lượng lớn hơn thì xác suất hộ là nghèo sẽ nhỏ hơn 50.

Giai đoạn 2: Áp dụng kết quả hồi quy đối với số liệu của Tổng điều tra

Trong giai đoạn 2 của phương pháp lập bản đồ nghèo đói, các hệ số hồi quy ước lượng từ bước 1 được kết hợp với số liệu của các đặc điểm tương tự của hộ từ cuộc Tổng điều tra để ước lượng xác suất mỗi hộ trong Tổng điều tra là nghèo. Điều này được thực hiện bằng cách cho các giá trị đặc điểm của hộ i từ Tổng điều tra, X_i^C vào phương trình 2. Xác suất kỳ vọng mà hộ i là nghèo được tính như sau:

$$E[P_i | X_i^C, \beta, \sigma^2] = \Phi \left[\frac{\ln z - X_i^C \beta}{\sigma} \right] \quad (3)$$

Sự ước lượng này không thực sự chính xác tuyệt đối cho từng hộ nhưng sẽ chính xác hơn cho nhiều hộ. Đối với từng khu vực nhất định (ví dụ như một huyện hay một tỉnh), Elbers và đồng nghiệp (2003) chỉ rằng rằng tỷ lệ dân số trong hộ là nghèo được ước lượng bằng trung bình của các xác suất mà từng hộ riêng biệt là nghèo:

$$E[P | X^C, \beta, \sigma^2] = \sum_{i=1}^N \frac{m_i}{M} \Phi \left[\frac{\ln z - X_i^C \beta}{\sigma} \right] \quad (4)$$

trong đó m_i là quy mô hộ i , M là tổng dân số của khu vực, N là số hộ, và X là ma trận của các đặc điểm hộ. Điểm mạnh của việc sử dụng số liệu của Tổng điều tra là cho phép ước tính tỷ lệ đói nghèo theo các đơn vị địa lý nhỏ hơn nhiều nếu chúng ta chỉ sử dụng VLSS.

Với điều kiện là a) sai số là đồng nhất, b) không có tự tương quan về không gian, và c) bộ số liệu Tổng điều tra được sử dụng, phương sai của tỷ lệ người đói nghèo (poverty headcount) ước lượng có thể tính như sau:

$$\text{var}(P^*) = \left(\frac{\partial P^*}{\partial \beta} \right)' \text{var}(\hat{\beta}) \frac{\partial P^*}{\partial \beta} + \left(\frac{\partial P^*}{\partial \sigma^2} \right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^N \frac{m_i^2 P_i^* (1-P_i^*)}{M^2} \quad (5)$$

trong đó n là số mẫu trong mô hình hồi quy. Do đó n , k , và σ^2 được lấy từ phân tích hồi quy, m_i , M , và N lấy từ Tổng điều tra dân số. Đạo hàm từng phần P^* với các tham số ước lượng có thể được tính như sau:

$$\frac{\partial P^*}{\partial \beta_j} = \sum_{i=1}^N \frac{m_i}{M} \left(\frac{-x_{ij}}{\hat{\sigma}} \right) \phi \left(\frac{\ln z - X_i^C \hat{\beta}}{\hat{\sigma}} \right) \quad (6)$$

$$\frac{\partial P^*}{\partial \hat{\sigma}^2} = -\frac{1}{2} \sum_{i=1}^N \frac{m_i}{M} \left(\frac{\ln z - X_i' \hat{\beta}}{\hat{\sigma}^3} \right) \phi \left(\frac{\ln z - X_i' \hat{\beta}}{\hat{\sigma}} \right) \quad (7)$$

Giá trị thứ nhất trong phương trình 5 biểu thị sai số mô hình sinh ra do thực tế là có một vài sự biến động (uncertainty) của giá trị đúng (true value) của β và σ trong phân tích hồi quy. Sự biến động này được xác định bằng ma trận đồng phương sai ước lượng (estimated covariance matrix of β) và phương sai ước lượng của σ^2 , cũng như tác động của sự thay đổi này tới P^* . Giá trị thứ 3 trong phương trình 5 đo lường sai số “idiosyncratic error” do thực tế là thậm chí là nếu β và σ được xác định chính xác thì các nhân tố cụ thể của hộ sẽ tạo ra giá trị chi tiêu thực tế khác với chi tiêu do ước lượng. Những phương trình này được miêu tả chi tiết hơn trong Hentschel và đồng nghiệp (2000) và Elbers và đồng nghiệp (2003).

Như đề cập ở trên, phương trình 5 chỉ đúng khi sử dụng bộ số liệu Tổng điều tra đầy đủ cho giai đoạn 2 của quá trình lập bản đồ đói nghèo. Trong nghiên cứu này, chúng tôi sử dụng mẫu 33% của Tổng điều tra trong giai đoạn 2, vì vậy phương trình 5 sẽ thay đổi như sau:

$$\text{var}(P^*) = \left(\frac{\partial P^*}{\partial \hat{\beta}} \right)' \text{var}(\hat{\beta}) \frac{\partial P^*}{\partial \hat{\beta}} + \left(\frac{\partial P^*}{\partial \hat{\sigma}^2} \right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^N \frac{m_i^2 P_i^* (1-P_i^*)}{M^2} + V_s \quad (8)$$

trong đó V_s là phương sai do sai số mẫu trong Tổng điều tra, do thiết kế mẫu. Trong nghiên cứu này, chúng tôi dựa trên phần mềm Stata để tính toán phương sai của sai số mẫu³.

Để so sánh tỷ lệ người đói nghèo ở các vùng hay tỉnh khác nhau, chúng ta có thể tính phương sai của sự khác nhau giữa hai tỷ lệ đói nghèo ước lượng được. Hentschel và các đồng nghiệp (2000, chú thích 17) cung cấp phương trình cho trường hợp khi bộ số liệu đầy đủ Tổng điều tra được sử dụng. Ở đây chúng tôi mở rộng phương trình gồm có phương sai do sai số mẫu:

$$\begin{aligned} \text{var}(P_1 - P_2) = & \left(\frac{\partial P_1 - P_2}{\partial \hat{\beta}} \right)' \text{var}(\hat{\beta}) \left(\frac{\partial P_1 - P_2}{\partial \hat{\beta}} \right) + \left(\frac{\partial P_1 - P_2}{\partial \hat{\sigma}^2} \right)^2 \frac{2\hat{\sigma}^4}{n-k-1} \\ & + V_i(P_1) + V_i(P_2) + V_s(P_1) + V_s(P_2) - 2 \text{cov}_s(P_1, P_2) \end{aligned} \quad (9)$$

trong đó $V_i(P_r)$ là phương sai idiosyncratic của tỷ lệ đói nghèo ước lượng cho vùng r (giá trị thứ 3 trong phương trình 5), $V_s(P_r)$ là phương sai mẫu của ước lượng đói nghèo cho vùng r và $\text{cov}_s(P_1, P_2)$ là đồng phương sai trong các ước lượng đói nghèo của vùng 1 và 2 do sai số mẫu.

³ điều này được tính bằng lệnh “svymean”. Stata tính các ước lượng tuyến tính (triển khai Taylor bậc 1) của phương sai mẫu dựa trên thông tin về sự phân tầng, đơn vị chọn mẫu sơ cấp, và các yếu tố quyền số. Xem Stata Corporation (2001, vol.4, svymean) để biết thêm các thông tin.

2.3 Phương pháp ước lượng các chỉ số đói nghèo khác

Các phương pháp ở trên cho phép ước lượng tỷ lệ người sống dưới mức đói nghèo. Chỉ tiêu này, thỉnh thoảng ký hiệu là P0 là chỉ tiêu trong tập các chỉ tiêu về đói nghèo của Foster, Greer, và Thorbecke (1984). Những chỉ tiêu về đói nghèo này có thể diễn tả như sau:

$$P_{\alpha} = \frac{1}{N} \sum_{i=1}^M \left[\frac{(z - y_i)}{z} \right]^{\alpha} \quad (10)$$

trong đó : z là chuẩn nghèo

y_i là thu nhập (hay chỉ tiêu của người i trong hộ nghèo

N là số người trong tổng thể,

M là số người nghèo

Các giá trị khác nhau của α trong phương trình 10 cho ta các chỉ tiêu nghèo đói khác nhau. Nếu $\alpha=0$, phương trình này cho ta tỷ lệ người đói nghèo. Vì giá trị trong ngoặc bằng 1, nên tổng của chúng cho ta số người trong các hộ nghèo, do đó khi chia cho N sẽ cho ta tỷ lệ người nghèo. Khi $\alpha=1$, công thức trên sẽ cho chúng ta chỉ tiêu khác, gọi là Khoảng cách nghèo. Khoảng cách nghèo không chỉ cho chúng ta biết bao nhiêu người nghèo mà còn cho chúng ta biết mức độ nghèo trung bình như thế nào. Chỉ tiêu này bằng tỷ lệ người nghèo (P_0) nhân với khoảng cách phần trăm trung bình giữa đường chuẩn nghèo và thu nhập của người nghèo. Khi $\alpha=2$, phương trình này cho ta chỉ tiêu khác, gọi là chỉ số khoảng cách nghèo bình phương. Chỉ số này cho biết không chỉ số người nghèo, nghèo như thế nào mà còn cho thấy mức độ bất bình đẳng về thu nhập giữa những người nghèo. Chỉ số này bằng P_0 nhân với bình phương khoảng cách giữa chuẩn nghèo và thu nhập của người nghèo.

Phương pháp lập bản đồ nghèo đói mô tả ở phần 3.1 và 3.3 cho ta một cách ước lượng tỷ lệ người nghèo dưới đường chuẩn nghèo, z , nhưng không cho ta thông tin về sự phân bố thu nhập giữa những người nghèo, do đó cần thiết phải tính P_1 và P_2 . Chúng tôi có thể áp dụng phương pháp đường chuẩn nghèo để ước lượng P_1 và P_2 bằng cách cho rằng z không phải là đường chuẩn nghèo. Đặc biệt, chúng ta có thể ước lượng sự phân bố của tổng thể theo chỉ tiêu bình quân đầu người bằng cách chạy lại các phép tính lập bản đồ nghèo đói cho các giá trị khác nhau của z . Cụ thể hơn, chúng tôi thực hiện các bước sau:

1. Chọn 100 mức chi tiêu bình quân đầu người khác nhau⁴, chia bằng nhau theo giá trị từ hộ giàu nhất tới hộ nghèo nhất
2. Đặt z bằng với mức thấp nhất trong 100 mức này (gọi là z_1), chạy các phép tính lập bản đồ đói nghèo để tính toán tỷ lệ dân số có mức chi tiêu bình quân dưới z_1
3. Tiếp theo, lặp lại bước 2 -3 đặt z bằng với 99 mức chi tiêu khác (z_2 tới z_{100}) giữ lại các giá trị z_i và tỷ lệ dân số dưới z_i trong một file cho phân tích tiếp theo.

Khi z tăng từ mức thấp nhất tới mức cao nhất, tỷ lệ dân số với chi tiêu bình quân đầu người dưới mức đó sẽ tăng từ 0 đến 100%. Vì vậy, các kết quả sẽ xác định cho ta phân phối tích lũy của dân số theo chi tiêu bình quân đầu người.

Thông tin này là đủ để tính toán các giá trị P_1 và P_2 . Khoảng cách của mỗi cặp z 's (z_i và z_{i+1}), chúng ta biết chi tiêu bình quân đầu người⁵ và tỷ lệ dân số với mức chi tiêu bình quân trong khoảng đó. Vì vậy, mỗi cặp z 's dưới đường đói nghèo có thể sử dụng đại diện cho một giá trị, có tính đến số hộ với chi tiêu bình quân đầu người trong khoảng đó.

2.4. Phương pháp đo lường bất bình đẳng

Trong nghiên cứu này, các chỉ số bất bình đẳng mô tả mức độ biến động của chi tiêu bình quân đầu người giữa các hộ. Sự bình đẳng tuyệt đối khi mà tất cả các hộ có mức chi tiêu bình quân đầu người như nhau, và sự bất bình đẳng tuyệt đối khi có tất cả mức chi tiêu khác hộ khác thì không có.

Trong phân tích này chúng tôi sẽ tính toán 3 chỉ số khá phổ biến đo lường sự bất bình đẳng: hệ số Gini, chỉ số Theil's L, và chỉ số Theil's T. Hai chỉ số sau là một phần trong một tập hợp của "chỉ số bất bình đẳng chung- geneal entropy", vì thế Theil L gọi là $GE(0)$ và Theil T gọi là $GE(1)$.

Hệ số Gini dựa trên đường cong Lorenze miêu tả phân bố tích lũy của thu nhập hay chi tiêu như là một hàm của phân bố tích lũy của hộ. Cụ thể hơn, hệ số Gini là phần diện tích trên đường cong Lorenz và dưới đường chéo 45% chỉ cho diện tích dưới đường chéo. Khi chúng ta có thông tin về tỷ lệ dân số dưới mức khác nhau của chi tiêu bình quân đầu người, hệ số Gini có thể ước tính như sau:

⁴ việc sử dụng 100 mức là tương đối tùy ý. Số mức lớn hơn sẽ cho mức độ chính xác cao hơn cho các ước lượng của phân phối tích lũy và các ước lượng của P_1 và P_2 . Tăng số mức, tất nhiên, cũng sẽ tăng các phép tính và thời gian chạy chương trình.

⁵ Nói một cách nghiêm túc, chúng tôi chỉ biết khoảng chi tiêu bình quân đầu người trong nhóm này và chúng tôi giả định rằng mức chung bình là $(z_i + z_{i+1})/2$. Nhưng nếu chúng tôi chọn số mức z lớn hơn, sự chênh lệch giữa z_i và z_{i+1} sẽ nhỏ hơn, vì vậy các sai số cho giả định này cũng sẽ nhỏ.

$$\text{Gini} = 2 \sum_{i=1}^N \left[\left(\frac{1}{2} (P_i + P_{i+1}) - \frac{1}{2} (X_i + X_{i+1}) \right) (P_{i+1} - P_i) \right] \quad (11)$$

trong đó P_i là tỷ lệ tích lũy của dân số trong khoảng i và X_i là tỷ lệ tích lũy của chi tiêu trong khoảng i . Giá trị đầu tiên trong ngoặc đơn lớn là độ cao của mỗi phần chia, từ đường chéo tới đường Lorenze, trong khi đó giá trị cuối cùng trong ngoặc đơn nhỏ là ‘độ rộng’ của mỗi phần chia. Hệ số Gini biến động từ 0 (bình đẳng tuyệt đối) tới 1 (bất bình đẳng tuyệt đối).

Chỉ số Theil L về bất bình đẳng được tính như sau:

$$\text{Theil L} = \text{GE}(0) = \frac{1}{N} \sum_{i=1}^N \ln \left(\frac{\bar{y}}{y_i} \right) \quad (12)$$

trong đó N là số hộ, $\bar{y}$ là chi tiêu bình quân đầu người trung bình, và y_i chi tiêu bình quân đầu người của hộ i . Chỉ số Theil L biến động từ 0 (bình đẳng tuyệt đối) và vô cùng (bất bình đẳng tuyệt đối).

Chỉ số bất bình đẳng Theil T được tính theo công thức sau::

$$\text{Theil T} = \text{GE}(1) = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\bar{y}} \ln \left(\frac{y_i}{\bar{y}} \right) \quad (13)$$

trong đó các biến được xác định như phương trình 12. Chỉ số Theil T biến động từ 0 (bình đẳng tuyệt đối) tới $\ln(N)$ (bất bình đẳng tuyệt đối).

Để có thể tính toán các chỉ số bất bình đẳng, chúng tôi sử dụng 5 bước miêu tả trong phần 2.4 để xây dựng phân bố tích lũy của các hộ theo chi tiêu bình quân đầu người. Để ước lượng hệ số Gini, chúng tôi tính phân bố tích lũy của chi tiêu từ giá trị của z_i và tỷ lệ của dân số ở các mức z . Sau đó, đối với mỗi giá trị của z , chúng tôi tính tổng của các chênh lệch giữa phân bố tích lũy của hộ (đường chéo trong đường Lorenz) và phân bố tích lũy của chi tiêu (đường Lorenz) nhân với tỷ lệ của dân số ở mức z đó. Tổng này chia cho 0.5 (diện tích dưới đường chéo) để có giá trị hệ số Gini.

Như cách tính P_1 và P_2 , hai chỉ số Theil được tính bằng cách sử dụng mỗi cặp z 's để biểu thị cho một giá trị của y_i . như miêu tả ở trên, giữa mỗi cặp của z 's (z_i và z_{i+1}), chúng ta biết được giá trị trung bình chi tiêu bình quân đầu người và tỷ lệ dân số với chi tiêu bình quân trong khoảng đó. thông tin này cho phép chúng ta áp dụng các phương trình 11 và 12 để tính các chỉ số Theil.

2.5 Hạn chế của phân tích này

Có hai hạn chế cần phải được đề cập trong việc thực hiện phương pháp lập bản đồ nghèo đói của Việt Nam. Thứ nhất, phân tích hồi quy trong Giai đoạn 1 không xem xét đến sự khác nhau phương sai của các biến phụ thuộc trong mẫu (heteroskedasticity). Mặt khác, bằng cách biểu thị biến phụ thuộc (chi tiêu bình quân đầu người) dưới dạng logarithm, chúng tôi có giảm mức độ không đồng đều về phương sai (heteroskedasticity). Các hệ số ước lượng từ phương pháp bình phương nhỏ nhất bình thường (ordinary least squares -OLS) cũng không bị chệch, nhưng các hệ số này cũng bị kém hiệu quả (“inefficient”) do đó, chúng ta cũng không sử dụng tất cả các thông tin có thể (xem StataCorp, 2001, Volume 4 “svyreg”).

Thứ hai, các hệ số ước lượng trong Giai đoạn 1 không xem xét đến trường hợp tự tương quan về không gian (spatial autocorrelation). Tự tương quan về không gian xuất hiện khi các biến phụ thuộc (hay sai số) của hồ quy trong một địa bàn của VLSS là tương quan với biến phụ thuộc (hay sai số) của địa bàn gần kề (vấn đề này sẽ thảo luận kỹ hơn ở trong Phần 2.6). Nếu các sai số có sự tương quan, các hệ số cũng sẽ không bị chệch nhưng kém hiệu quả. Điều này có thể xảy ra nếu một số các yếu tố khác (như khoảng cách tới thành phố chính) không có trong mô hình hồi quy và bị tự tương quan về mặt không gian. Ví dụ, tất cả các hộ gần một thành phố có thể có sai số âm (chi tiêu ước lượng thấp hơn chi tiêu thực tế). Nói cách khác, nếu chi tiêu bình quân đầu người trong một địa bàn bị ảnh hưởng của chi tiêu bình quân đầu người ở địa bàn bên cạnh trong mẫu VLSS, thì các hệ số hồi quy ước lượng sẽ bị chệch. Với các khoảng cách giữa các địa bàn kề nhau trong VLSS, điều này có thể ít xảy ra.

Các nhà nghiên cứu của Ngân hàng Thế giới gần đây có nghiên cứu về các vấn đề phương sai không đồng đều và tự tương quan về mặt không gian trong Giai đoạn 1. Những giải pháp phân tích cho sự khác nhau về tỷ lệ người đói nghèo là không thể trong trường hợp này, và nó trở nên rất cần thiết cho việc sử dụng phương pháp mô phỏng phức tạp hơn để tính các ước lượng và sai số chuẩn (xem Elbers và đồng nghiệp, 2003). Mặc dù phân tích ban đầu cho thấy sự hiện diện của một số tự tương quan về mặt không gian, nhưng chúng tôi cũng không thể loại bỏ hoàn toàn điều này bằng cách cho một số biến cấp xã vào trong phân tích hồi quy. Điều này cho thấy, có thể có một số kết quả không hoàn toàn chính xác trong phân tích hồi quy của Giai đoạn 1, dù những tác động này rất khó đánh giá.

2.6 Phương pháp phân tích hồi quy không gian tổng thể (global spatial regression analysis)

Phân tích hồi quy về không gian (spatial regression analysis) thực hiện trong dự án này liên quan đến ước lượng đói nghèo là một hàm của các biến về nông nghiệp, khí hậu và tiếp cận thị trường. Phần này sẽ miêu tả Phân tích hồi quy về không gian toàn bộ, trong đó “toàn bộ” muốn nói đến thực tế rằng mô hình giả định quan hệ giữa đói nghèo và các yếu tố về địa lý là như nhau trong giữa các vùng trong cả nước. Biến phụ thuộc là tỷ lệ đói nghèo huyện ước tính từ phân tích bản đồ đói nghèo miêu tả ở phần 3. Biến độc lập được liệt kê ở Bảng 2.

Bảng 2. Các biến giả thích sử dụng trong phân tích hồi quy về không gian

Các biến ngoại sinh	Các biến nội sinh có thể
% diện tích ở các độ cao khác nhau	Dân số, mật độ dân số
% diện tích ở các độ dốc khác nhau	% dân số ở thành thị
địa hình	số và loại thị trường
Loại đất	Mật độ đường xá
loại đất mặt	Mật độ sông
lượng mưa	thời gian vận chuyển giữa các khu vực
nhiệt độ	
số giờ nắng	
khoảng cách tới các thành phố	

Như thảo luận ở Phần 2.5, một vấn đề liên quan tới việc phân tích hồi quy dựa trên quan hệ về không gian là sự tự tương quan của số liệu⁶. Nhìn chung, tự tương quan về không gian có nghĩa là các biến trong một khu vực có thể bị ảnh hưởng bởi giá trị của biến đó trong khu vực lân cận. Có 2 cách mà vấn đề này có thể xảy ra:

Sự phụ thuộc trễ về không gian (Spatial lag dependence) đề cập đến trường hợp biến phụ thuộc trong 1 vùng bị tác động bởi biến phụ thuộc ở các vùng bên cạnh. Ví dụ, nếu biến phụ thuộc là thu nhập hay đói nghèo, thì có thể mức độ các hoạt động kinh tế trong một vùng bị ảnh hưởng trực tiếp bởi mức độ các hoạt động kinh tế ở vùng lân cận thông qua sự di dân, buôn bán, hay liên kết về đầu tư. Mô hình phụ thuộc trễ về không gian có thể miêu tả dưới dạng sau:

⁶ Một vấn đề liên quan là các số liệu về địa lý phải tổng hợp thành cho các đơn vị hành chính và các kết quả có thể bị ảnh hưởng tùy theo cách mà chúng ta thực hiện. Phụ Lục A sẽ miêu tả vấn đề này cụ thể hơn.

$$y_i = \sigma \sum_{j \neq i} w_{ij} y_j + X_j \beta + \varepsilon_j \quad (14)$$

trong đó y_i là biến phụ thuộc cho vùng i ,

σ là hệ số tự tương quan không gian,

w_{ij} là quyền số không gian phản ánh mức độ gần kề của i và j ,

y_j là biến phụ thuộc cho vùng j ,

X_j là ma trận biến giải thích,

β là vector các hệ số, và

ε là sai số.

Ma trận các quyền số không gian, w , mô tả mức độ gần kề giữa từng đôi quan sát theo không gian. Thường nó là biến nhị nguyên tùy theo liệu hai vùng có gần kề nhau không hay một biến gần kề tùy theo một vài hàm của khoảng cách giữa hai vùng. Nếu phân tích hồi quy được thực hiện không có sự điều chỉnh cho sự phụ thuộc trễ về không gian, thì các hệ số ước lượng có thể bị lệch (biased) và không đồng nhất.

Loại thứ 2 của vấn đề này có thể xảy ra là sự phụ thuộc về sai số không gian, trong đó sai số ở vùng này tự tương quan với sai số vùng bên cạnh. Điều này có thể xảy ra nếu các biến không trong mô hình hồi quy nhưng có ảnh hưởng tới biến phụ thuộc và chúng sẽ tự tương quan về không gian. Ví dụ, trình độ của chính quyền địa phương tác động tới thu nhập và đói nghèo nhưng nó rất khó để có thể cho vào trong mô hình. Vì trình độ của chính quyền địa phương thì có thể bị tự tương quan về không gian (tất cả các thành phố ở một bang có thể bị ảnh hưởng bởi trình độ của chính quyền bang đó), sai số trong mỗi vùng có thể bị tương quan với các vùng bên cạnh. Mô hình này có thể viết dưới dạng như sau:

$$y_i = X_j \beta + \lambda \sum_{j \neq i} w_{ij} y_j \varepsilon_j + \varepsilon_i \quad (15)$$

trong đó y_i là biến phụ thuộc cho vùng i ,

λ là hệ số tự tương quan về không gian,

w_{ij} là quyền số không gian phản ánh mức độ gần kề của i và j ,

X_j là ma trận biến giải thích,

β là vector của các hệ số, và

ε là sai số.

Trong trường hợp này, việc sử dụng bình phương nhỏ nhất (ordinary least squares) để ước lượng mô hình không nhận được các hệ số lệch (biased coefficients), nhưng những ước lượng của hệ số sẽ không hiệu quả và kiểm định t và F sẽ tạo ra những kết luận không chính xác.

Trong nghiên cứu này, chúng tôi ước lượng tỷ lệ đói nghèo cấp huyện (P0) là hàm của các biến không gian liệt kê ở Bảng 2. Kiểm định Chow cho thấy các hệ số ước tính tỷ lệ đói nghèo ở thành thị khác nhau đáng kể với các hệ số ước tính tỷ lệ đói nghèo nông thôn. Vì thế, chúng tôi thực hiện phân tích hồi quy riêng biệt cho nông thôn và thành thị.

Ma trận quyền số được tạo ra bằng cách sử dụng nghịch đảo khoảng cách giữa các trung tâm của hai huyện. Nói cách khác, giá trị của w_{ij} bằng nghịch đảo khoảng cách giữa trung tâm của huyện i trung tâm của huyện j.

Kiểm định hệ số nhân tử Lagrange được sử dụng để kiểm định ý nghĩa thống kê của σ và λ , cho thấy sự cần thiết để sử dụng mô hình trễ phụ thuộc về không gian (spatial dependence lag model) và mô hình phụ thuộc sai số về không gian (spatial error dependence model). Thường với những mô hình hồi quy không gian, cả hai tham số là có ý nghĩa thống kê và cách hay dùng là áp dụng mô hình cho những giá trị hệ số nhân tử Lagrange cao hơn.

Phân tích này được thực hiện bằng cách sử dụng module “spatreg” viết cho Stata có trên trang web Stata (www.stata.com).

2.7 Các phương pháp trong phân tích hồi quy theo vùng

Mô hình tổng thể (global model) được miêu tả trong Phần 2.6 giả định rằng quan hệ giữa đói nghèo và các yếu tố địa lý là như nhau giữa các vùng trong cả nước. Phân tích hồi quy không gian theo vùng (local spatial regression analysis) không giả định như vậy và kiểm tra sự khác nhau về không gian trong tương quan giữa đói nghèo và các yếu tố địa lý. Chúng tôi sử dụng một khung phân tích hồi quy của sổ chuyển động (moving window regression framework) trong đó các mô hình hồi quy khác nhau sẽ được ước lượng, mỗi trung tâm trên “điểm hồi quy” (“regression point”) và bao gồm các quan sát kề bên được xác định bởi “đai bao bọc hạt nhân” (“kernel bandwidth.”). Các ước lượng hệ số của các vùng được tạo ra cho mỗi điểm hồi quy.

Một mô hình dựa trên kỹ thuật hồi quy quyền số địa lý (GWR) (Brunsdon, và cộng sự 1996), ở đó các quan sát giữa các cửa sổ hồi quy vùng được tính theo khoảng cách tới các điểm hồi quy sẽ được áp dụng. Các quan sát gần với điểm hồi quy x_i có được trọng số cao hơn so với các quan sát ở xa hơn. Cửa sổ hồi quy quyền số sau đó được dịch chuyển tới điểm hồi quy tiếp theo đến khi tất cả các điểm được nằm vào trong đó.

Vì phương pháp này dựa trên khung hồi quy cổ truyền, nên kỹ thuật sẽ tạo ra kết quả hồi quy chuẩn cho từng điểm hồi quy. Điều này cho phép kết kết quả hồi quy (gồm các hệ số và R^2)

có thể vẽ trên bản đồ, cho thấy sự khác nhau giữa các vùng. điều này làm cho kỹ thuật đó đặc biệt hữu ích đối với việc phân tích quan hệ giữa các dữ liệu về không gian.

Một mô hình hồi quy toàn bộ chuẩn được viết dưới dạng sau:

$$y_i = a_0 + \sum_j x_{ij} a_j + \varepsilon \quad (16)$$

có thể mở rộng cho mô hình hồi quy vùng như sau:

$$y_i = a_0 + \sum_j x_{ij} a_{ij} + \varepsilon \quad (17)$$

trong đó y là biến phụ thuộc,

x là biến độc lập,

a là hệ số hồi quy ,

i là chỉ số cho vùng (location),

j là chỉ số cho biến độc lập, và

ε là sai số

Cho mỗi hồi quy tại các điểm i, quyền số của các quan sát được lựa chọn phụ thuộc vào khoảng cách tới điểm hồi quy j. Hàm phân tách khoảng cách Gauss áp dụng trong phân tích này có thể viết như sau :

$$w_{ij} = \exp\left[-\frac{1}{2}\left(d_{ij}/b\right)^2\right] \quad (18)$$

trong đó w là quyền số,

d là khoảng cách từ các điểm hồi quy ,

b là phạm vi hay bán kính của sự ảnh hưởng xung quanh mỗi quan sát

Bên cạnh đó, chúng ta có thể kiểm định liệu mô hình vùng có thực sự biểu thị tương quan tốt hơn mô hình toàn bộ bằng cách so sánh giá trị R^2 của 2 mô hình. Hơn nữa, Fotheringham và cộng sự (2002) cho rằng kiểm định Monte Carlo sự khác nhau về không gian trong các tham số ước lượng là có ý nghĩa thống kê. Những kiểm định tham gia điều chỉnh ngẫu nhiên các địa điểm của các quan sát theo thời gian, chạy GWR và sau đó so sánh thống kê các ước lượng tham số của sự phân bố địa lý thực tế.

3. Đói nghèo và bất bình đẳng xét về mặt không gian

3.1 Đặc điểm của hộ có liên quan đến chỉ tiêu bình quân đầu người

Như miêu tả ở trên, bước đầu tiên trong việc xây dựng bản đồ nghèo đói là ước lượng chỉ tiêu bình quân đầu người là hàm của các biến có trong cả Tổng điều tra Dân số và Điều tra mức sống dân cư Việt Nam. Những đặc điểm của hộ bao gồm quy mô hộ, thành phần, dân tộc, trình độ chủ hộ (hay vợ/chồng chủ hộ), nghề nghiệp chủ hộ, loại nhà, tiếp cận tới các dịch vụ cơ bản, sở hữu một số đồ dùng. Bảng 3 liệt kê các biến sử dụng mô tả đặc điểm của hộ trong phân tích hồi quy. Trong mỗi tập hợp biến mô tả như phân loại vùng thì một vùng sẽ bị bỏ qua và xem như nhóm tham khảo. Ví dụ, Miền núi phía Bắc là nhóm bị bỏ qua cho biến “region”, vì thế các hệ số vùng khác phản ánh tác động của mức sống của vùng đó so với vùng núi phía Bắc.

Cũng hợp lý nếu chúng ta kỳ vọng rằng các hệ số để ước lượng chỉ tiêu trong khu vực nông thôn có thể khác với các hệ số ước lượng chỉ tiêu ở khu vực thành thị. Các kiểm định thống kê cho thấy rằng các hệ số trong mô hình phân tích thành thị là tương đối khác trong mô hình phân tích vùng nông thôn⁷. Điều này cho thấy nên có sự phân tích riêng biệt cho mẫu thành thị và nông thôn.

Trong một phân tích trước đây, chúng tôi cố gắng ước lượng mô hình cho 2 vùng thành thị và 7 vùng nông thôn (xem Minot và Baulch, 2002a). Các kết quả hồi quy không làm chúng tôi hài lòng với giá trị R^2 thấp, nhiều hệ số không có ý nghĩa về thống kê và một số hệ số còn sai dấu. Vì vậy chúng tôi đã áp dụng những mô hình thành thị-nông thôn trong phân tích này.

Những kết quả của phân tích hồi quy được trình bày ở Bảng 4. Cả mô hình thành thị, nông thôn đều có thể giải thích được trên 50% của sự biến động chỉ tiêu bình quân đầu người. Đây chỉ là một kết quả tương đối tốt khi phân tích số liệu cross-section, nhưng nó rất hữu ích vì các yếu tố khác là không có trong mô hình giải thích một nửa của sự biến động .

Theo những kết quả của Bảng 4, những hộ quy mô lớn thường có chỉ tiêu bình quân đầu người thấp đối với cả thành thị và nông thôn. Dấu âm của các hệ số về quy mô hộ cho thấy, khi các yếu tố khác không đổi, mỗi thành viên thêm sẽ làm giảm 7-8% của chỉ tiêu bình quân đầu người.

⁷ Kiểm định Chow loại bỏ giả thiết rằng các hệ số cho mẫu thành thị là giống mẫu cho nông thôn ($F=6.16, p<.001$).

Ở khu vực nông thôn những hộ có tỷ lệ người già, trẻ em hay phụ nữ cao hơn thì thường nghèo hơn. Tuy nhiên, ở thành thị tỷ lệ phụ nữ nhiều hơn lại không quan trọng, nhưng tỷ lệ trẻ em cao hơn thì vẫn tương quan chặt chẽ với đói nghèo. Thành phần của hộ có ảnh hưởng ít hơn đối với các hộ thành thị bởi vì khả năng kiếm thu nhập ở thành phố là ít phụ thuộc vào thể chất hơn.

Bảng 3. Mô tả các biến sử dụng trong mô hình

Biến	Loại biến	Mô tả biến
lnrpce	Liên tục	Log chi tiêu bình quân đầu người
hhszise	Liên tục	số thành viên hộ
pelderly	Liên tục	Tỷ lệ người trên 60 tuổi (%)
pchild	Liên tục	tỷ lệ thành viên dưới 15 tuổi (%)
pfemale	Liên tục	Tỷ lệ nữ (%)
ethnic	nhị nguyên	Chủ hộ là dân tộc thiểu số
Iedchd_1	nhị nguyên	chủ hộ chưa tốt nghiệp tiểu học (loại bỏ)
Iedchd_2	nhị nguyên	chủ hộ tốt nghiệp tiểu học
Iedchd_3	nhị nguyên	Chủ hộ tốt nghiệp phổ thông cơ sở
Iedchd_4	nhị nguyên	Chủ hộ tốt nghiệp phổ thông trung học
Iedchd_5	nhị nguyên	Chủ hộ tốt nghiệp trung cấp-kỹ thuật
Iedchd_6	nhị nguyên	Chủ hộ có trình độ sau trung học
Iedcsp_0	nhị nguyên	chủ hộ có vợ/chồng
Iedcsp_1	nhị nguyên	chủ hộ có vợ/chồng chưa tốt nghiệp tiểu học (loại bỏ)
Iedcsp_2	nhị nguyên	Vợ/chồng chủ hộ tốt nghiệp tiểu học
Iedcsp_3	nhị nguyên	Vợ/chồng Chủ hộ tốt nghiệp phổ thông cơ sở
Iedcsp_4	nhị nguyên	Vợ/chồng Chủ hộ tốt nghiệp phổ thông trung học
Iedcsp_5	nhị nguyên	Vợ/chồng Chủ hộ tốt nghiệp trung cấp-kỹ thuật
Iedcsp_6	nhị nguyên	Vợ/chồng Chủ hộ có trình độ sau trung học
Ioccup_1	nhị nguyên	Chủ hộ là quan chức hay quản lý
Ioccup_2	nhị nguyên	Chủ hộ là cán bộ
Ioccup_3	nhị nguyên	Chủ hộ là người bán hàng hay làm dịch vụ
Ioccup_4	nhị nguyên	Chủ hộ làm nông/lâm ngư nghiệp
Ioccup_5	nhị nguyên	Chủ hộ là công nhân có tay nghề cao
Ioccup_6	nhị nguyên	Chủ hộ là công nhân phổ thông
Ioccup_7	nhị nguyên	Chủ hộ không làm việc (bỏ)
Ihouse_1	nhị nguyên	Nhà kiên cố
Ihouse_2	nhị nguyên	Nhà bán kiên cố
Ihouse_3	nhị nguyên	Nhà tạm thời
htypla1	Liên tục	Tương tác của log(diện tích nhà) và Ihouse_1
htypla2	Liên tục	Tương tác của log(diện tích nhà) và Ihouse_2

electric	nhị nguyên	Nhà có điện
Inwate_1	nhị nguyên	Nhà sử dụng nước từ vòi nước công cộng hay cá nhân
Inwate_2	nhị nguyên	Nhà sử dụng nước giếng
Inwate_3	nhị nguyên	Nhà sử dụng nước sông hay hồ (loại bỏ)
Itoile_1	nhị nguyên	Nhà có toilet tự hoại
Itoile_2	nhị nguyên	Nhà có toilet hai ngăn
Itoile_3	nhị nguyên	Không có toilet tự hoại hay hai ngăn (bỏ)
tv	nhị nguyên	Hộ có TV
radio	nhị nguyên	Hộ có đài
reg7_1	nhị nguyên	Hộ ở miền núi phía Bắc (bỏ)
reg7_2	nhị nguyên	Hộ ở ĐBSH
reg7_3	nhị nguyên	Hộ ở Bắc Trung Bộ
reg7_4	nhị nguyên	Hộ ở Nam Trung Bộ
reg7_5	nhị nguyên	Hộ ở Tây Nguyên
reg7_6	nhị nguyên	Hộ ở Đông Nam Bộ
reg7_7	nhị nguyên	Hộ ở ĐBSCL
Nguồn:	Điều tra Mức sống dân cư Việt Nam 1997-98	

Dân tộc thiểu số⁸ là là yếu tố đáng ngạc nhiên vì có tác động ít tới chi tiêu bình quân đầu người khi các yếu tố khác không đổi. Ở nông thôn, các hệ số về dân tộc thiểu số chỉ có ý nghĩa thống kê 10%, trong khi ở thành thị thì không có ý nghĩa thống kê. Điều này không có nghĩa là dân tộc thiểu số là có thu nhập bằng dân tộc khác. Mà nó có nghĩa là sau khi đề cập các chỉ số về quy mô, thành phần hộ, trình độ giáo dục, đặc điểm nhà cửa, sở hữu tài sản và các đặc điểm khác, dân tộc thiểu số không đóng góp nhiều thêm trong việc ước lượng chi tiêu bình quân đầu người (xem Bảng 4).

Ở cả khu vực và thành thị, trình độ học vấn của chủ hộ (Iedchd_) là yếu tố tốt để ước lượng chi tiêu bình quân đầu người (khi 'không đi học' là biến bị loại bỏ). Có 5 biến biểu thị trình độ của chủ hộ và đều có ý nghĩa thống kê ở mức 1% cho cả thành thị và nông thôn (xem Bảng 5). Ở thành thị, các hộ có chủ hộ tốt nghiệp tiểu học hay phổ thông cơ sở dường như không giàu hơn các hộ có chủ hộ chưa tốt nghiệp tiểu học, những chủ hộ có trình độ ở mức cao hơn thì thu nhập cũng cao hơn nhiều.

Nhìn chung, trình độ giáo dục của vợ/chồng chủ hộ không có ảnh hưởng nhiều như trình độ của chủ hộ khi ước lượng chi tiêu bình quân đầu người. Ở nông thôn, chỉ 2 cấp học vấn cao nhất của vợ/chồng chủ hộ (trung cấp trở lên và sau trung học) cho thấy sự khác nhau với mức trình độ tham chiếu (chưa tốt nghiệp tiểu học). Trình độ của vợ/chồng chủ hộ lại có ý nghĩa hơn đối với hộ thành thị (xem Bảng 4). Điều này cho thấy là trình độ của vợ/chồng chủ hộ có

⁸ Dân tộc thiểu số được định nghĩa là tất cả các dân tộc thiểu số trừ dân tộc Kinh và Hoa

thể cho lợi ích khác mà không chỉ đo lường bằng chỉ tiêu bình quân đầu người, ví dụ như sức khoẻ và dinh dưỡng của trẻ em tốt hơn.

Nghề nghiệp của chủ hộ cũng có ý nghĩa quan trọng trong ước lượng chỉ tiêu bình quân đầu người ở thành thị và nông thôn (xem Bảng 4 và Bảng 5). Ở nông thôn, 3 nhóm ngành đầu tiên (quan chức hay quản lý, cán bộ hay kỹ thuật, và bán hàng hay dịch vụ) là có thu nhập cao hơn so với chủ hộ không có việc làm (biến loại bỏ). Sự khác nhau về chỉ tiêu bình quân đầu người của hộ với chủ nông nghiệp và phi nông nghiệp là không có ý nghĩa thống kê. Một kết quả ngạc nhiên là chủ hộ không có việc làm bao gồm người nghỉ hưu cũng như tỷ lệ mất cân đối của lao động có tay nghề có thể cố gắng tìm việc làm cho thu nhập cao hơn.

Ở thành thị, những hộ mà chủ hộ là quan chức/nhà quản lý là có thu nhập cao hơn nhiều so với hộ mà chủ hộ không có việc làm, trong khi những hộ có chủ hộ là lao động có tay nghề kém lại có thu nhập thấp nhất. Điều này cho thấy rằng ở khu vực thành thị, chủ hộ không có việc làm là nhân tố không đáng tin cậy để cho rằng hộ đó là nghèo (xem Bảng 4)..

Bảng 4. Mô hình hồi quy log của chỉ tiêu bình quân đầu người

	Mô hình nông thôn		Mô hình thành thị		
N	4269		1730		
R-bình phương	0.536		0.550		
Biến	hệ số	t	biến	Hệ số	t
hhsz	-0.0772	-19.5 ***	hhsz	-0.0785	-8.1 ***
pelderly	-0.0831	-2.4 **	pelderly	-0.1026	-1.6
pchild	-0.3353	-9.4 ***	pchild	-0.2368	-3.6 ***
pfemale	-0.1177	-3.5 ***	pfemale	0.0386	0.5
ethnic	-0.0765	-1.9 *	ethnic	0.0142	0.2
ledchd_2	0.0585	3.4 ***	ledchd_2	0.0616	1.7
ledchd_3	0.0883	4.5 ***	ledchd_3	0.0338	1.3
ledchd_4	0.0884	3.3 ***	ledchd_4	0.1368	3.2 ***
ledchd_5	0.1355	4.2 ***	ledchd_5	0.1603	3.5 ***
ledchd_6	0.2552	4.9 ***	ledchd_6	0.1843	3.7 ***
ledcsp_0	0.0173	1.0	ledcsp_0	0.0344	0.8
ledcsp_2	0.0049	0.3	ledcsp_2	0.0642	1.9 *
ledcsp_3	0.0132	0.6	ledcsp_3	0.0987	2.6 **
ledcsp_4	0.0107	0.3	ledcsp_4	0.1912	2.7 **
ledcsp_5	0.0921	2.3 **	ledcsp_5	0.1285	3.2 ***
ledcsp_6	0.1571	2.7 ***	ledcsp_6	0.1752	3.1 ***
loccup_1	0.1414	3.5 ***	loccup_1	0.2312	3.0 ***
loccup_2	0.1350	3.3 ***	loccup_2	0.0576	1.2
loccup_3	0.1362	3.4 ***	loccup_3	0.0357	0.9
loccup_4	-0.0163	-0.6	loccup_4	-0.0093	-0.2

Ioccup_5	0.0701	1.9 *	Ioccup_5	0.0071	0.2
Ioccup_6	-0.0586	-1.7 *	Ioccup_6	-0.1599	-2.9 ***
Ihouse_1	-0.9228	-4.3 ***	Ihouse_1	-0.5194	-3.4 ***
Ihouse_2	-0.3120	-3.6 ***	Ihouse_2	-0.4001	-3.8 ***
htypla1	0.2958	5.7 ***	htypla1	0.2001	5.4 ***
htypla2	0.1180	5.2 ***	htypla2	0.1403	4.6 ***
electric	0.0765	2.7 ***	electric	-0.0026	0.0
Inwate_1	0.0828	1.4	Inwate_1	0.2289	5.3 ***
Inwate_2	0.1157	4.4 ***	Inwate_2	0.0340	0.6
Itoile_1	0.2700	5.5 ***	Itoile_1	0.1311	2.2 **
Itoile_2	0.0556	2.6 **	Itoile_2	0.0049	0.1
tv	0.2124	15.1 ***	tv	0.2167	5.5 ***
radio	0.1009	7.0 ***	radio	0.1599	6.2 ***
Ireg7_2	0.0314	0.6	Ireg7_2	0.0693	0.7
Ireg7_3	0.0485	0.8	Ireg7_3	0.0445	0.6
Ireg7_4	0.1373	2.2 **	Ireg7_4	0.1460	1.9 *
Ireg7_5	0.1708	2.1 **	Ireg7_5	Omitted	
Ireg7_6	0.5424	9.4 ***	Ireg7_6	0.4151	5.5 ***
Ireg7_7	0.3011	5.1 ***	Ireg7_7	0.1895	2.1 **
_cons	7.5327	108.7 ***	_cons	7.7538	64.7 ***

Nguồn: Phân tích hồi quy Điều tra mức sống dân cư 1997-98.

Chú ý: * hệ số có ý nghĩa ở mức 10 %, ** ở mức 5 %, và *** mức 1%. VLSS không có hộ ở

Bảng 5. Ý nghĩa thống kê của các nhóm biến

Khu vực	Biến	Df1	Df2	giá trị F	Gía trị P	
Nông thôn	Trình độ chủ hộ	5	129	7.80	0.0000	***
	Trình độ vợ/chồng chủ hộ	6	129	1.97	0.0738	*
	Nghề nghiệp chủ hộ	6	129	12.65	0.0000	***
	Loại nhà	2	129	14.00	0.0000	***
	nguồn nước chính	2	129	9.69	0.0001	***
	Loại nhà vệ sinh	2	129	15.64	0.0000	***
	Vùng	6	129	26.20	0.0000	***
Thành thị	Trình độ chủ hộ	5	55	4.01	0.0036	***
	Trình độ vợ/chồng chủ hộ	6	55	3.10	0.0110	**
	Nghề nghiệp chủ hộ	6	55	2.90	0.0157	**
	Loại nhà	2	55	10.76	0.0001	***
	nguồn nước chính	2	55	17.17	0.0000	***
	Loại nhà vệ sinh	2	55	4.12	0.0216	**
	Vùng	5	55	10.29	0.0000	***

Nguồn: Phân tích hồi quy Điều tra mức sống dân cư 1997-98.

Chú ý: Biến độc lập là log chi tiêu bình quân đầu người

*hệ số có ý nghĩa ở mức 10 %, ** ở mức 5 %, và *** mức 1%

Các đặc điểm khác nhau của nhà ở là yếu tố tốt để giải thích cho chi tiêu bình quân đầu người. Các hộ có nhà kiên cố có chi tiêu bình quân đầu người cao hơn 19% so với hộ có nhà tạm thời ở nông thôn và 24% so với hộ có nhà tạm thời ở khu vực thành thị. Tương tự, hộ có nhà bán kiên cố thì có chi tiêu bình quân cao hơn hộ có nhà tạm thời. Diện tích nhà ở cũng là một yếu tố tốt để đánh mức sống của hộ. Hộ ở Việt Nam có diện tích trung bình 45 mét vuông và cứ 10% tăng lên của diện tích thì tương ứng với sự tăng lên từ 12–30% chi tiêu bình quân đầu người, tùy theo vị trí của hộ (thành thị, nông thôn) hay loại nhà (kiên cố hay bán kiên cố).ⁱ

Điện⁹ cũng là yếu tố có ý nghĩa thống kê trong phân tích chi tiêu của hộ ở nông thôn nơi có tới 71% hộ có điện. Ngược lại ở thành thị, nơi có tới 98% hộ có điện thì điện lại không phải yếu tố có ảnh hưởng nhiều tới chi tiêu của hộ (xem Bảng 4).

Nguồn nước chính cũng có ý nghĩa phân biệt hộ nghèo và không nghèo. Trong khu vực nông thôn, các hộ có sử dụng nước giếng có chi tiêu bình quân cao hơn các hộ sử dụng nước sông hay hồ (biến loại bỏ). Tiếp cận tới nước không phải là yếu tố có ý nghĩa thống kê tác động tới

⁹ Cụ thể hơn, biến này muốn đề cập đến các loại chiếu sáng sử dụng bởi hộ gia đình.

chi tiêu ở nông thôn, có thể bởi vì chỉ 2% số hộ là nằm trong nhóm này. Ngược lại, trong khu vực thành thị, hơn một nửa số mẫu (58%) có thể tiếp cận tới nguồn nước máy, và biến này là yếu tố tốt để giải thích chi tiêu bình quân đầu người ở thành thị.

Nhà vệ sinh cũng có thể xem là yếu tố phân biệt hộ nghèo và không nghèo. Ở khu vực nông thôn, toilet tự hoại và 2 ngăn là có ý nghĩa thống kê cho hộ có chi tiêu bình quân cao hơn ở mức ý nghĩa 5%. Ở thành thị, hộ có toilet tự hoại là yếu tố có ý nghĩa thống kê ở mức 5%, nhưng toilet 2 ngăn thì không có ý nghĩa thống kê khi phân tích chi tiêu bình quân đầu người (xem Bảng 4).

Sở hữu TV là một trong yếu tố có ý nghĩa nhất tác động tới chi tiêu bình quân đầu người và có ý nghĩa thống kê đối với hộ thành thị và nông thôn. Đài cũng là yếu tố tốt tác động chi tiêu bình quân đầu người, có ý nghĩa thống kê ở mức 1% đối với các hộ thành thị và nông thôn. Như kỳ vọng, hệ số cho sở hữu đài nhỏ hơn hệ số cho sở hữu TV.

Các biến giả vùng có trong cả mô hình thành thị và nông thôn, với miền núi phía Bắc là vùng tham chiếu (base region). Ngay cả khi các đặc điểm khác của hộ không thay đổi, hộ nông thôn ở 4 vùng miền Nam cho kết quả giàu hơn hộ nông thôn ở các vùng miền núi phía Bắc. Hệ số của vùng Đông nam Bộ là cao nhất, chi tiêu bình quân của hộ ở vùng này cao hơn 72% hộ ở miền núi phía Bắc. Một xu hướng tương tự cho các hộ ở thành thị (xem Bảng 4). Các biến giả vùng đều có ý nghĩa thống kê ở mức 1% đối với cả thành thị và nông thôn (xem Bảng 5)

3.2 Tỷ lệ đói nghèo

Tỷ lệ đói nghèo được định nghĩa là tỷ lệ dân số sống trong các hộ gia đình có chi tiêu bình quân đầu người dưới đường đói nghèo chung theo định nghĩa của GSO (2000, trang 260). Đây là chỉ số Foster-Greer-Thorbecke về đói nghèo khi $\alpha=0$, được ký hiệu là P_0 . Chúng tôi sẽ lần lượt trình bày ước lượng tỷ lệ đói nghèo cấp xã, huyện và tỉnh.

Tỷ lệ đói nghèo toàn quốc

Tỷ lệ đói nghèo trung bình trên toàn quốc được ước lượng bằng cách áp dụng phương pháp ước lượng diện tích nhỏ sử dụng mẫu 33% của Tổng điều tra dân số và nhà ở 1999 là khoảng 35.9%. So với tỷ lệ đói nghèo ước theo VLSS1998 là 37,4% (xem Nhóm Nghiên cứu về đói nghèo, 1999) và các ước lượng khu vực nhỏ trước đây, dựa trên mẫu 3% của Tổng điều tra dân số và nhà ở 1999 là 36,5% (xem Minot và Baulch, 2002a.). Đây không phải là ước lượng đói nghèo độc lập vì phương pháp ước lượng diện tích nhỏ dựa trên số liệu VLSS 1998 trong Giai đoạn 1, nhưng không khẳng định rằng các đặc điểm của hộ gia đình trong Tổng điều tra dân số có giá trị giống nhau so với những đặc điểm ở VLSS 1998.

Tỷ lệ đói nghèo tỉnh (P_0)

Bảng 6 cho chúng ta thấy ước lượng tỷ lệ đói nghèo, trong khi đó Hình 1 cho ta thấy sự phân bố đói nghèo về mặt không gian. Trong bản đồ, vùng nghèo nhất có màu da cam đậm và vùng ít nghèo nhất có màu xanh thẫm. Những kết quả phân tích càng khẳng định rằng đói nghèo ở Đông Bắc và Tây Bắc là cao nhất, đặc biệt là ở các tỉnh giáp biên giới với Trung Quốc và biên giới Tây Bắc giáp với Lào. Cụ thể hơn, tỷ lệ đói nghèo cao nhất (70-80 %) ở các tỉnh Lai Châu, Hà Giang và Sơn La. Màu cam nhạt cho biết tỷ lệ đói nghèo ở mức là 60-70 % ở Lào Cai, Cao Bằng, Lạng Sơn, và Bắc Kạn. Tỷ lệ đói nghèo thấp hơn nhưng vẫn ở mức trên 50% ở các tỉnh màu vàng, các tỉnh bên trong vùng Đông Bắc và Tây Bắc (Hòa Bình, Tuyên Quang, và Yên Bái), phía Bắc của Tây Nguyên (Gia Lai và Kon Tum), và các tỉnh ven biển miền Trung (Ninh Thuận và Quảng Trị). Một điều rất đáng chú ý là tất cả 10 tỉnh nghèo nhất ở Tây Bắc và Đông Bắc.

Bảng 6. Tỷ lệ đói nghèo (P_0) thành thị và nông thôn theo tỉnh

Mã	Tỉnh	Xếp hạng (1=nghèo nhất)	Chung		Nông thôn		Thành thị	
			Tỷ lệ đói nghèo (P_0)	Sai số chuẩn	Tỷ lệ đói nghèo (P_0)	Sai số chuẩn	Tỷ lệ đói nghèo (P_0)	Sai số chuẩn
101	Ha Noi	55	0.16	0.013	0.31	0.026	0.05	0.011
103	Hai Phong	51	0.29	0.020	0.40	0.029	0.08	0.017
105	Ha Tay	33	0.39	0.027	0.41	0.029	0.13	0.024
107	Hai Duong	47	0.33	0.026	0.36	0.030	0.12	0.026
109	Hung Yen	39	0.37	0.027	0.39	0.029	0.17	0.036
111	Ha Nam	35	0.38	0.028	0.40	0.030	0.14	0.028
113	Nam Dinh	41	0.35	0.026	0.38	0.030	0.11	0.024
115	Thai Binh	43	0.34	0.029	0.36	0.030	0.08	0.020
117	Ninh Binh	35	0.38	0.025	0.42	0.029	0.10	0.023
201	Ha Giang	2	0.75	0.020	0.81	0.022	0.24	0.033
203	Cao Bang	5	0.67	0.024	0.75	0.027	0.17	0.041
205	Lao Cai	4	0.70	0.018	0.79	0.021	0.21	0.027
207	Bac Kan	7	0.60	0.027	0.67	0.031	0.21	0.037
209	Lang Son	6	0.62	0.024	0.73	0.028	0.17	0.033
211	Tuyen Quang	9	0.57	0.030	0.61	0.033	0.14	0.023
213	Yen Bai	9	0.57	0.025	0.67	0.031	0.17	0.025
215	Thai Nguyen	25	0.43	0.033	0.50	0.042	0.16	0.022
217	Phu Tho	20	0.45	0.038	0.50	0.044	0.15	0.022
219	Vinh Phuc	20	0.45	0.044	0.48	0.049	0.20	0.028
221	Bac Giang	17	0.46	0.042	0.48	0.046	0.16	0.024
223	Bac Ninh	35	0.38	0.043	0.40	0.048	0.18	0.026
225	Quang Ninh	41	0.35	0.025	0.51	0.041	0.15	0.024

301	Lai Chau	1	0.80	0.014	0.88	0.015	0.19	0.029
303	Son La	3	0.73	0.020	0.81	0.022	0.14	0.023
305	Hoa Binh	8	0.59	0.028	0.65	0.032	0.14	0.023
401	Thanh Hoa	17	0.46	0.034	0.49	0.037	0.14	0.025
403	Nghe An	17	0.46	0.034	0.50	0.037	0.12	0.022
405	Ha Tinh	20	0.45	0.036	0.48	0.040	0.15	0.028
407	Quang Binh	15	0.47	0.033	0.51	0.038	0.14	0.026
409	Quang Tri	13	0.51	0.027	0.59	0.034	0.22	0.031
411	Thua Thien - Hue	15	0.47	0.026	0.58	0.036	0.20	0.027
501	Da Nang	55	0.16	0.017	0.34	0.032	0.11	0.020
503	Quang Nam	29	0.41	0.029	0.46	0.034	0.17	0.029
505	Quang Ngai	20	0.45	0.030	0.49	0.034	0.15	0.027
507	Binh Dinh	35	0.38	0.028	0.45	0.036	0.16	0.028
509	Phu Yen	29	0.41	0.029	0.46	0.035	0.18	0.031
511	Khanh Hoa	47	0.33	0.022	0.44	0.032	0.14	0.022
601	Kon Tum	13	0.51	0.037	0.65	0.052	0.20	0.028
603	Gia Lai	11	0.53	0.037	0.63	0.049	0.20	0.027
605	Dac Lac	25	0.43	0.045	0.50	0.056	0.17	0.025
701	TP Ho Chi Minh	61	0.05	0.008	0.08	0.012	0.05	0.009
703	Lam Dong	43	0.34	0.035	0.46	0.055	0.14	0.021
705	Ninh Thuan	11	0.53	0.026	0.63	0.032	0.21	0.032
707	Binh Phuoc	54	0.17	0.020	0.19	0.023	0.08	0.017
709	Tay Ninh	57	0.13	0.016	0.14	0.018	0.08	0.017
711	Binh Duong	60	0.08	0.008	0.08	0.011	0.06	0.012
713	Dong Nai	58	0.11	0.012	0.13	0.016	0.06	0.010
715	Binh Thuan	20	0.45	0.026	0.53	0.034	0.24	0.035
717	Ba Ria-Vung Tau	59	0.10	0.010	0.13	0.016	0.06	0.010
801	Long An	51	0.29	0.022	0.32	0.026	0.14	0.026
803	Dong Thap	33	0.39	0.025	0.42	0.028	0.19	0.032
805	An Giang	31	0.40	0.022	0.46	0.027	0.21	0.032
807	Tien Giang	53	0.27	0.024	0.30	0.027	0.12	0.024
809	Vinh Long	47	0.33	0.024	0.36	0.027	0.15	0.027
811	Ben Tre	50	0.32	0.024	0.34	0.026	0.15	0.028
813	Kien Giang	31	0.40	0.023	0.45	0.027	0.21	0.034
815	Can Tho	43	0.34	0.022	0.40	0.027	0.14	0.026
817	Tra Vinh	25	0.43	0.026	0.47	0.030	0.19	0.030
819	Soc Trang	25	0.43	0.025	0.48	0.029	0.22	0.033
821	Bac Lieu	40	0.36	0.023	0.41	0.028	0.20	0.034
823	Ca Mau	43	0.34	0.024	0.38	0.028	0.17	0.028

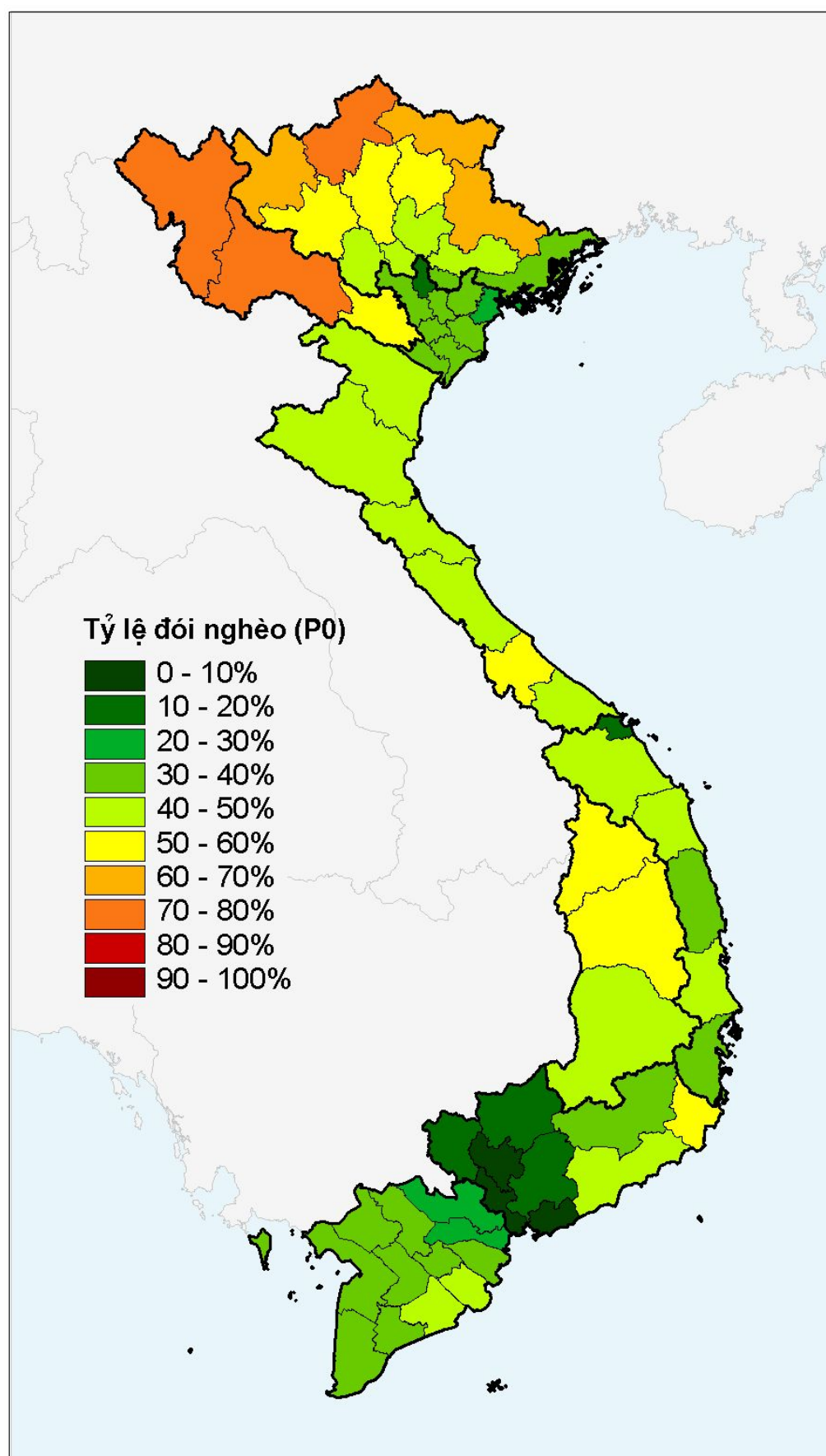
Nguồn: Phân tích VLSS 1997-98 và Tổng điều tra dân số nhà ở 1999 .

Những tỉnh có tỷ lệ nghèo đói thấp nhất là những tỉnh gần các trung tâm thành phố lớn. Ba tỉnh miền Nam (Hồ Chí Minh, Bình Dương, và Bà Rịa-Vũng Tàu) là màu xanh đậm, có nghĩa là tỷ lệ đói nghèo khoảng 10% hay thấp hơn. Các tỉnh khác, bao gồm Đồng Nai, Tây Ninh, Đà Nẵng, Hà Nội, và Bình Phước có tỷ lệ đói nghèo ước lượng từ 10-20%.

Hầu hết các tỉnh ở ĐBSH, ĐBSCL và Duyên hải miền Trung có màu xanh đậm, tỷ lệ đói nghèo từ 20-50%.

Bản đồ nghèo đói là rất có ích trong việc xác định đói nghèo về mặt không gian, cho chúng ta số liệu cụ thể hơn, bao gồm cả sai số chuẩn của các ước lượng đói nghèo và tỷ lệ đói nghèo thành thị và nông thôn cho mỗi tỉnh. Một trong những điểm mạnh của phương pháp lập bản đồ nghèo đói là tính được các sai số chuẩn, đo lường mức độ chính xác của các ước lượng (Hộp 1 giải thích rõ hơn về sai số chuẩn).

Hình 1. Bản đồ tỷ lệ đói nghèo (P_0) theo tỉnh



Hộp 1. Giải thích sai số chuẩn và khoảng tin cậy

Cũng như các phương pháp đo lường đói nghèo, phương pháp lập bản đồ nghèo đói cũng không thể cho các kết quả hoàn toàn chính xác. Trong Giai đoạn 1, những đặc điểm của hộ không thể ước tính giá trị chỉ tiêu một cách hoàn hảo được. Thậm chí ngay cả khi điều đó có thể thực hiện được, thì vẫn có sự khác nhau giữa các hộ trong VLSS và các hộ trong tổng điều tra dân số. Hơn thế nữa, vì mẫu phân tích Tổng điều tra của chúng ta chỉ có 33%, nên sẽ có sai số mẫu.

Một số yếu tố ảnh hưởng tới mức độ chính xác của các ước lượng về nghèo đói. Thứ nhất, nếu như các phương trình hồi quy trong giai đoạn 1 là rất tốt để ước lượng chỉ tiêu hộ gia đình dựa trên các đặc điểm của hộ, thì các ước lượng đói nghèo sẽ chính xác hơn. Thứ hai, độ chính xác của các ước lượng nghèo đói có xu hướng cao hơn cho những vùng có tỷ lệ đói nghèo xấp xỉ 0% hay gần 100%. Thứ ba, mức độ chính xác là cao hơn cho những vùng có số nhiều hộ giống nhau hơn là những vùng với chỉ vài hộ giống nhau và vùng mà các hộ có sự khác biệt.

Sai số chuẩn giúp chúng ta xác định độ lớn của sai số xung quanh những tỷ lệ đói nghèo ước lượng. Có 95% cơ hội ước lượng đói nghèo là “đúng” nằm trong khoảng 2 sai số chuẩn của ước lượng nghèo đói. Ví dụ, trường hợp của tỉnh Yên Bái, tỷ lệ đói nghèo ước tính là 57% và sai số là 0.025. Điều này có nghĩa là 95% khoảng tin cậy các ước lượng đói nghèo 57% là $\pm 5\%$ (0.025×2). Nói cách khác 95% cơ hội tỷ lệ đói nghèo đúng của Yên Bái là nằm trong khoảng từ 52% và 62%.

Các tỷ lệ đói nghèo và các khoảng tin cậy cho 61 tỉnh (theo thứ tự từ ít nghèo nhất cho đến tỷ lệ đói nghèo cao nhất) được trình bày ở Hình 2. Những điểm hình kim cương là tỷ lệ đói nghèo tỉnh ước tính, còn những đường ngang là những giới hạn phía trên và dưới của khoảng tin cậy 95%. Đồ thị này chỉ ra rằng hầu hết các tỉnh (36) có tỷ lệ đói nghèo từ 20% đến 50%. Ngược lại, có một số tỉnh có tỷ lệ đói nghèo dưới 20% hoặc trên 50%, và khoảng cách giữa những tỷ lệ đói nghèo của các tỉnh gần nhau là tương đối lớn..

Giữa các tỉnh, khoảng tin cậy 95 % (xem Hộp 1) từ $\pm 1.6\%$ đến $\pm 9\%$, với khoảng tin cậy trung bình ± 5.2 . Một nửa số tỉnh có khoảng tin cậy từ ± 4.2 và ± 5.8 ¹⁰. Đắk Lắk có khoảng tin cậy cao nhất ($\pm 9\%$). Có điều không rõ ràng là tại sao sai chuẩn lại tương đối lớn đối với Đắk Lắk, có thể các hộ có thu nhập từ cả phê sử dụng tiền vào các mục đích trong trong 17 đặc điểm của hộ mà chúng ta phân tích.

Một ý nghĩa rất quan trọng của sai số chuẩn là nếu 2 tỉnh mà khác nhau 5%, chẳng hạn như trường hợp sự khác nhau này không có ý nghĩa về mặt thống kê. Ví dụ, nếu tỉnh A có tỷ lệ nghèo đói 40% và tỉnh B có tỷ lệ nghèo đói là 44 %, nhìn chung chúng ta không thể nói rằng

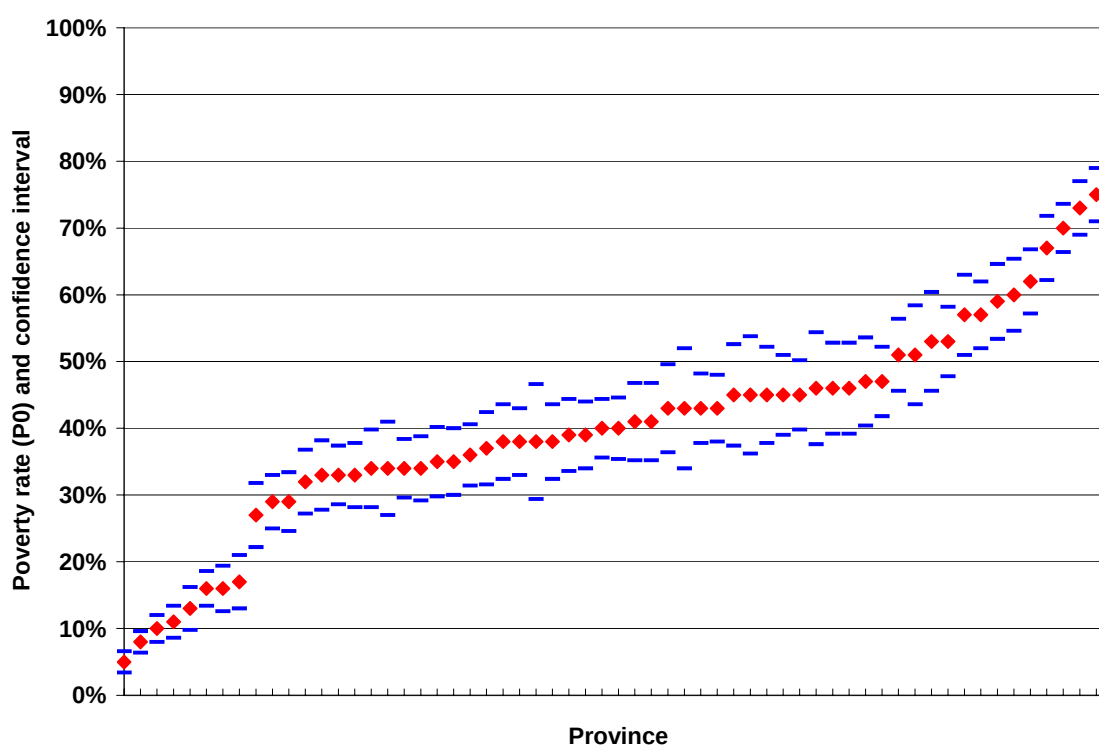
¹⁰ Cụ thể hơn, điều này đề cập đến khoảng 4 bậc (inter-quartile range). Giá trị đầu tiên (± 4.2) là nhóm 25% và giá trị thứ hai (± 5.8) là nhóm 75%. Vì thế, hơn một nửa số tỉnh nằm trong khoảng này.

tỉnh B là nghèo hơn tỉnh A. Theo quy luật chung, hai tỷ lệ đói nghèo phải khác nhau ít nhất 8-10% để cho chúng ta sự tin cậy là sự khác nhau này là có ý nghĩa về mặt thống kê.

Và chúng ta có thể nhìn vào Hình 2, khoảng tin cậy là nhỏ hơn rất nhiều khi tỷ lệ đói nghèo là khá thấp hay khá cao. Khi tỷ lệ đói nghèo dưới 20% hay trên 70%, khoảng tin cậy sẽ là ± 2 hay ± 3 %. Ngược lại, khi tỷ lệ đói nghèo chỉ khoảng 50%, khoảng tin cậy có thể sẽ lên tới 5% hay cao hơn.

Sự khác nhau về tỷ lệ nghèo đói giữa nông thôn và thành thị của mỗi tỉnh cũng rất có ý nghĩa (xem Bảng 6). Trong 61 tỉnh tỷ lệ đói nghèo của khu vực nông thôn đều cao hơn khu vực thành thị. Thực chất, trong khi tỷ lệ đói nghèo ở nông thôn biến động lớn từ 8% cho đến gần 90%, thì tỷ lệ đói nghèo ở tất cả thành thị là dưới 25%.

Hình 2. Đồ thị tỷ lệ đói nghèo theo tỉnh và khoảng tin cậy



Tỷ lệ đói nghèo cấp huyện (P_0)

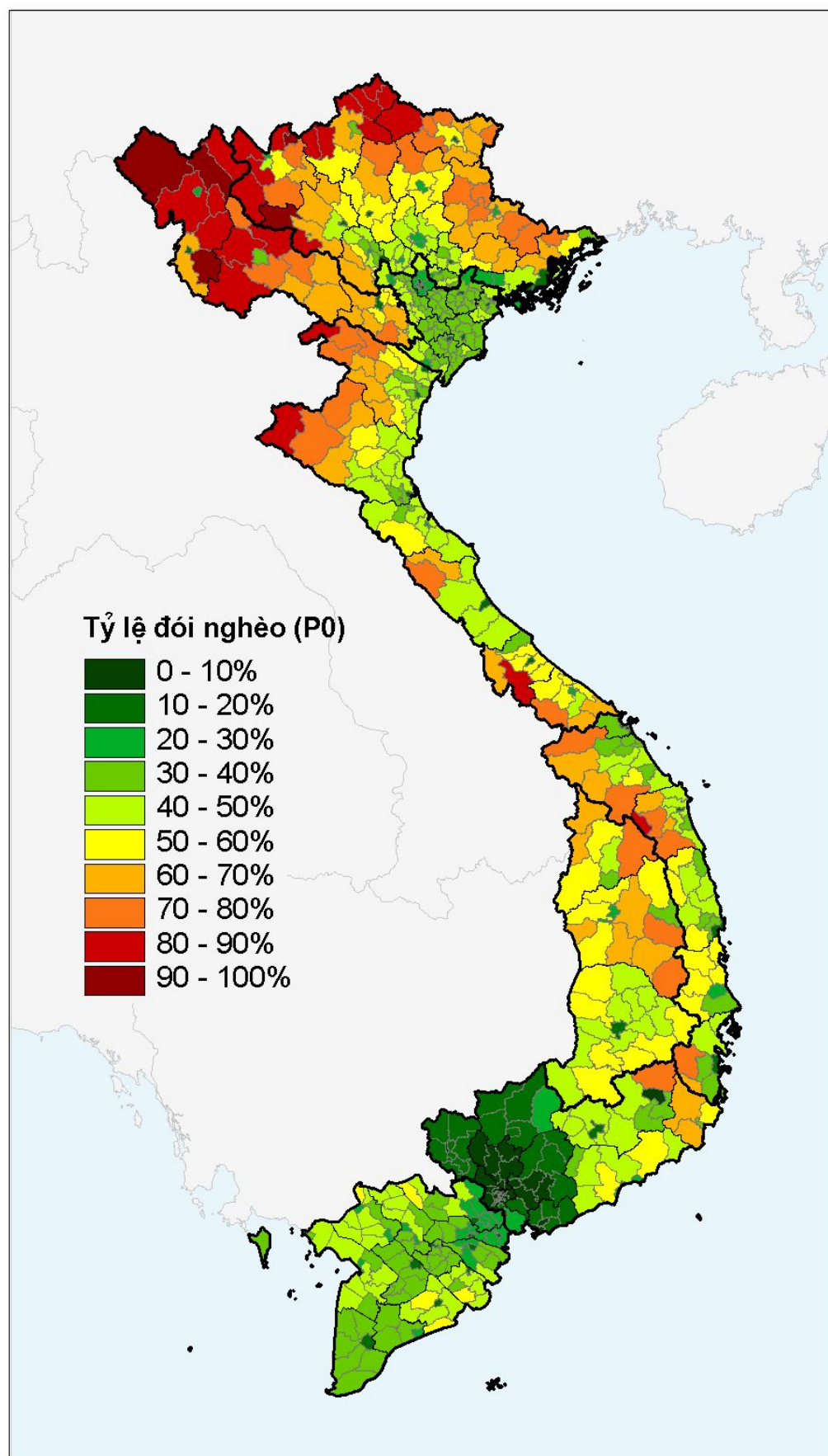
Phương pháp lập bản đồ nghèo đói có thể sử dụng để tính toán các chỉ số đói nghèo cấp huyện. Tỷ lệ đói nghèo cấp huyện có thể nhìn thấy trong Hình 3. Bản đồ đói nghèo cấp huyện cho ta thấy cụ thể hơn nhiều so với bản đồ đói nghèo cấp tỉnh. Ví dụ như trong bản đồ đói nghèo tỉnh, Sơn La, Lai Châu có thể thấy màu cam, biểu thị tỷ lệ đói nghèo trong khoảng từ 70-80%. Tuy nhiên bản đồ đói nghèo cấp huyện cho thấy sự khác nhau rõ ràng về đói nghèo giữa hai tỉnh, tỷ lệ đói nghèo dưới 70% ở các huyện đông nam Sơn La (các huyện Yên Châu,

Phú Yên, và Mộc Châu) và trên 90 % ở vùng xa góc Tây bắc của Lai Châu (Mông Tề và Sín Hồ).

Tương tự, bản đồ tỉnh cho thấy hầu hết các tỉnh Bắc Trung Bộ là có màu xanh nhạt, biểu thị tỷ lệ đói nghèo trong khoảng 40-50 %. Ngược lại, bản đồ huyện cho thấy tỷ lệ đói nghèo của các huyện dọc theo duyên hải là dưới 40% nhưng ở các tỉnh bên trong thì một vài huyện có tỷ lệ nghèo đói cao hơn 70%. Hai huyện trong vùng có tỷ lệ đói nghèo trên 80%, cả hai giáp biên giới Lào: huyện Mường Lat ở Thanh Hoá và Kỳ Sơn ở Nghệ An. Bản đồ đói nghèo cấp huyện cũng cho ta thấy sự khác nhau về tình trạng đói nghèo giữa các tỉnh Tây Nguyên và ĐBSCL, tuy nhiên ở bản đồ đói nghèo cấp tỉnh thì không thấy được.

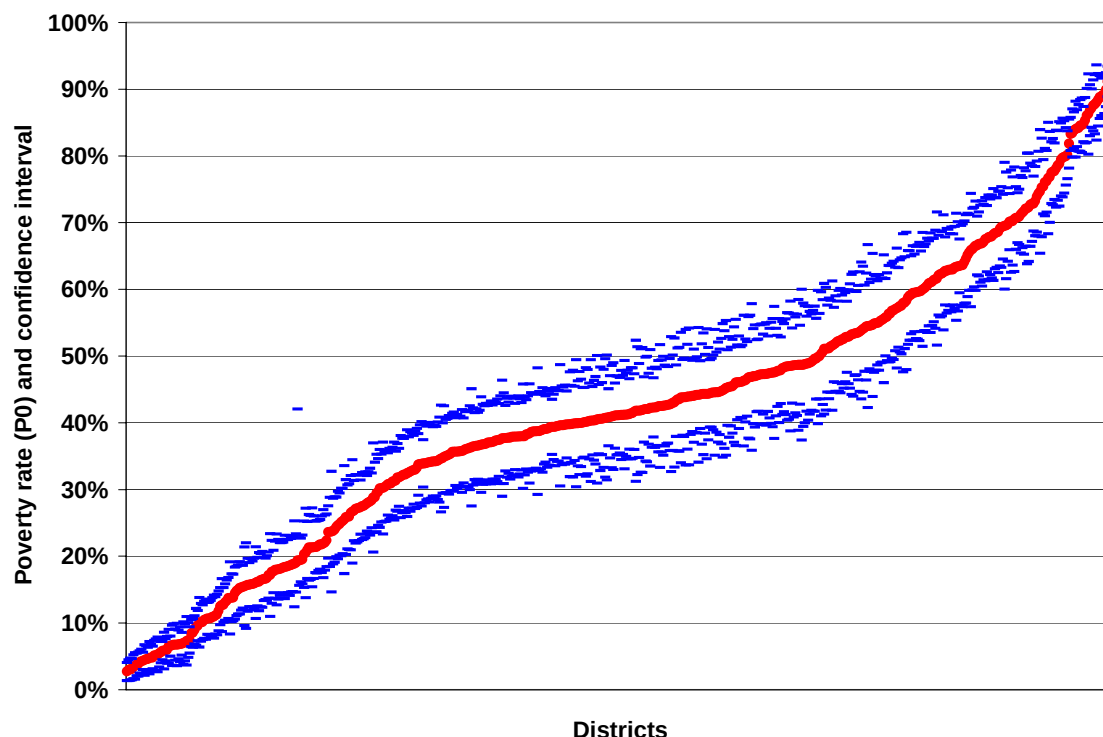
Cũng cần lưu ý là, sự chênh lệch rõ ràng giữa tỷ lệ đói nghèo thấp ở Đông nam Bộ (năm 1998) và tỷ lệ cao hơn ở Tây Nguyên và Nam trung Bộ phần nào do việc sử dụng các biến giả vùng trong Giai đoạn 1

Hình 3. Bản đồ đói nghèo cấp huyện (P₀)



Khoảng tin cậy 95% cho tỷ lệ đói nghèo cấp huyện được trình bày trong Hình 4. Như trong Hình 2, đường nằm ở giữa là các ước lượng của tỷ lệ nghèo đói, còn hai đường dọc theo ở trên và ở dưới chính là giới hạn trên và dưới của khoảng tin cậy 95%.

Hình 4. Tỷ lệ đói nghèo huyện và khoảng tin cậy

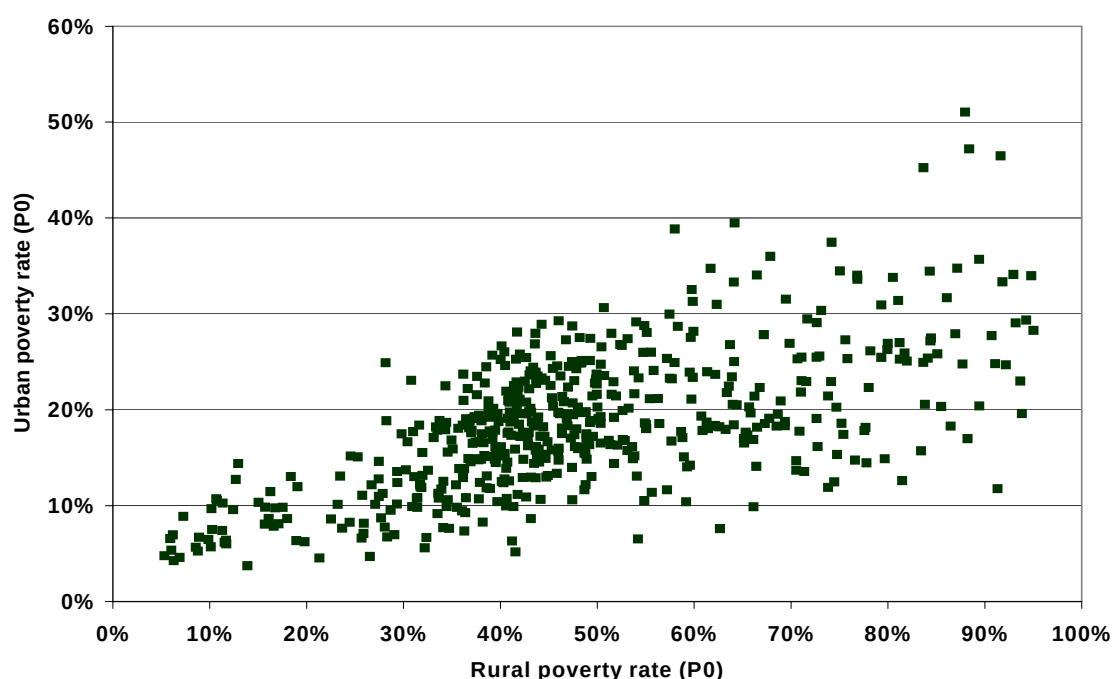


Khoảng tin cậy của tỷ lệ đói nghèo huyện dao động từ ± 1.3 đến ± 22 %, với mức trung bình là ± 5.8 %. Một nửa số huyện có khoảng tin cậy trong khoảng ± 4.4 và ± 6.9 %. Như đề cập ở trước, khoảng tin cậy là nhỏ hơn (và các ước lượng đói nghèo chính xác hơn) khi tỷ lệ đói nghèo là xấp xỉ bằng 0 hay 100%. Khi tỷ lệ đói nghèo nằm ở khoảng giữa (40-50%), khoảng tin cậy có xu hướng trong khoảng từ ± 5 đến ± 10 %.

Nhìn chung, các khoảng tin cậy của tỷ lệ đói nghèo huyện lớn hơn tỷ lệ đói nghèo tỉnh, trung bình là ± 5.2 %. Tỉnh có tỷ lệ đói nghèo thiếu chính xác nhất là Bạch Long Vĩ (tỉnh Hải Phòng): tỷ lệ đói nghèo ước lượng là $19\% \pm 22$ %. Nguyên nhân của sự thiếu chính xác này là ước lượng trên chỉ dựa vào 18 hộ trên đảo. Tuy nhiên đây là một trường hợp không thể chấp nhận được. Khoảng tin cậy lớn thứ hai là $\pm 12\%$. Hơn nữa, chỉ có 6 trong tổng số 614 huyện có ít hơn 1000 hộ trong mẫu của tổng điều tra, trong khi đó, 90% số đó có trên 2500 hộ.

Như đề cập ở trên, tỷ lệ đói nghèo của thành thị thấp hơn tỷ lệ đói nghèo ở khu vực nông thôn. Hầu hết tỷ lệ đói nghèo ở thành thị của các huyện trong khoảng 10-30%, trong khi của khu vực nông thôn từ dưới 10% cho đến trên 90%, với phần lớn các huyện trong khoảng từ 30-60 %. Có điều khá thú vị là, có một sự tương quan cùng chiều giữa tỷ lệ đói nghèo thành thị và tỷ lệ đói nghèo của nông thôn trong cùng một huyện (xem Hình 5).

Hình 5. Tỷ lệ đói nghèo của thành thị và nông thôn cấp huyện



Tỷ lệ đói nghèo cấp xã (P_0)

Phương pháp lập bản đồ nghèo đói cũng có thể ước lượng tỷ lệ đói nghèo cấp xã (P_0). Tuy nhiên cũng cần phải thận trọng khi sử dụng các kết quả này bởi vì số lượng hộ trong một xã rất nhỏ vì vậy độ tin cậy của các tỷ lệ đói nghèo ước lượng cho các xã này sẽ không cao. Độ tin cậy về tỷ lệ đói nghèo của các xã sẽ được thảo luận hơn nữa trong phần sau.

Tỷ lệ đói nghèo của các xã được trình bày trong Hình 6. Bản đồ cho chúng ta thấy chi tiết hơn so với bản đồ cấp huyện (xem Hình 3). Ví dụ, trong vùng Tây Bắc và Đông Bắc, bản đồ cấp xã cho biết số chấm xanh trong vùng da cam và vùng đỏ. Đây là những vùng thành thị với tỷ lệ đói nghèo tương đối thấp, xung quanh là những vùng nông thôn với tỷ lệ đói nghèo cao hơn rất nhiều. Bên cạnh đó, đường đi của Sông Hồng có thể nhận thấy là đường màu vàng và màu da cam khi vào Việt Nam từ Tây Bắc và xuống thẳng Hà Nội. Các xã gần sông Hồng có lợi thế là đất đai bằng phẳng, thủy lợi và vận chuyển đường thủy, tất cả các yếu tố này giúp giảm tỷ lệ đói nghèo.

Trong vùng Đông Bắc, sự ảnh hưởng của mạng lưới đường xá là rất rõ ràng ở một số vùng. Ví dụ, có một đường màu vàng da cam mở tới Đông nam từ biên giới Trung Quốc qua Cao Bằng và Lạng Sơn. Đây là đường cao tốc (Quốc Lộ 4A và 4B) đi từ biên giới Trung Quốc qua trung tâm của 2 tỉnh ra biển. Điều này có thể phản ánh tác động của việc tiếp cận thị trường tới đói nghèo, hay nó cũng có thể là những con đường này được xây ở những vùng ít núi đồi có tiềm năng nông nghiệp lớn (so sánh Hình 6 và Hình 7).

Thành phố Hà Nội với những chấm màu xanh đậm, xung quanh là màu xanh và vàng nhạt.

điều này cho thấy, các xã ở Hà Nội có tỷ lệ đói nghèo dưới 10%, so với tỷ lệ từ 20-40 % ở hầu hết các xã xung quanh của ĐBSH.

Ở Bắc Trung Bộ, bản đồ đói nghèo cấp xã thậm chí còn mô phỏng rõ hơn, các xã ven biển có tỷ lệ đói nghèo thấp hơn trong khi các xã trong vùng lục địa dọc biên giới Lào có tỷ lệ đói nghèo cao hơn rất nhiều. Chỉ có một ngoại trừ là vùng phía Tây của thành phố Vinh. Tỷ lệ đói nghèo thấp của khu vực này có thể giải thích là do ảnh hưởng của của đường quốc lộ (Quốc lộ 8) chạy từ phía tây Vinh tới biên giới Lào, đây là kênh quan trọng thương mại giữa hai nước. Một con đường khác (Quốc lộ 7) chạy từ tây bắc của Vinh tới biên giới Lào ở Kỳ Sơn (xem Hình 6 và Hình 7).

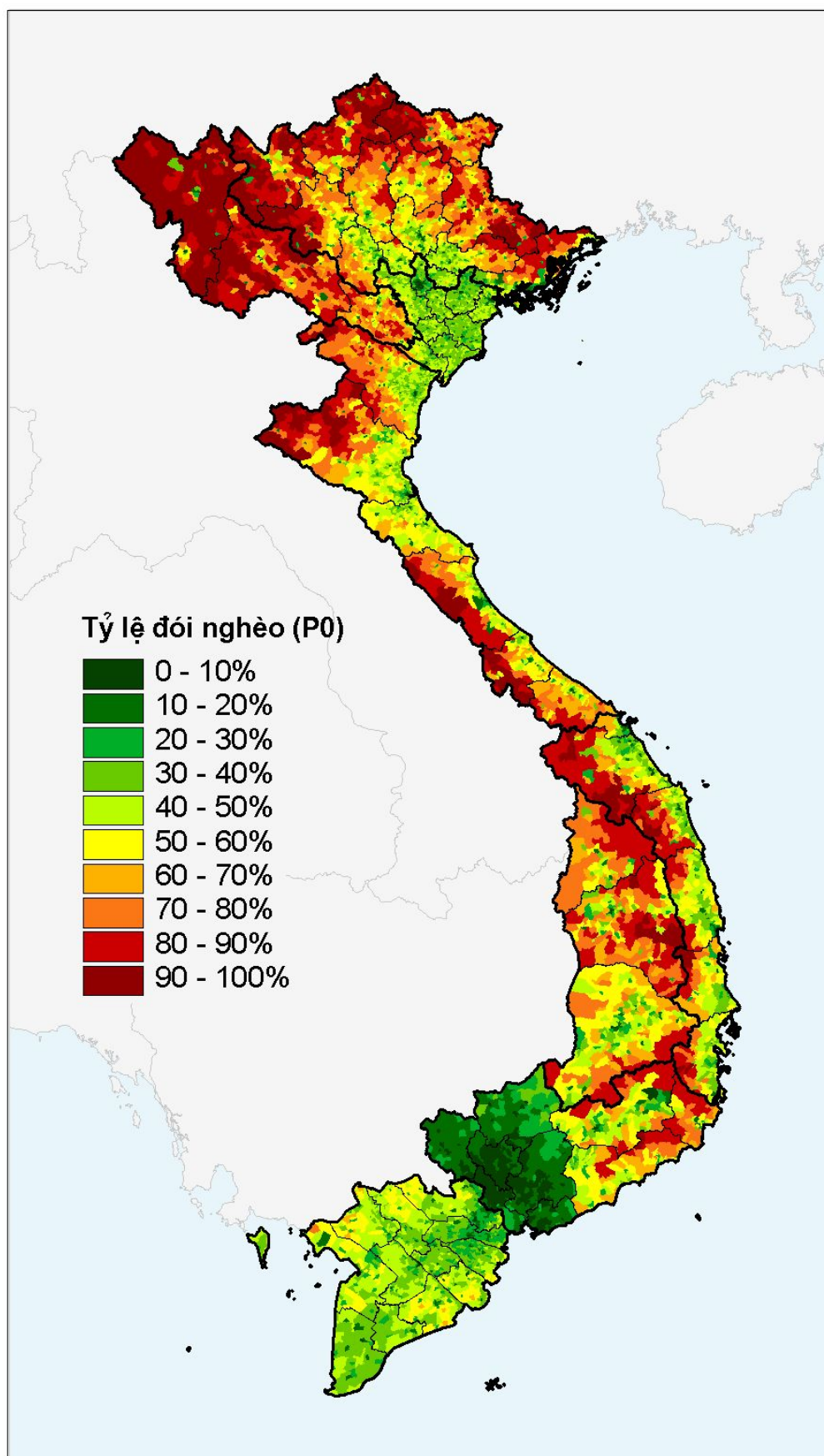
Duyên hải Nam trung Bộ là một trong những vùng nghèo nhất của Việt Nam. Nhất là các tỉnh vùng ven biển Bình Thuận và Linh Thuận với màu da cam trên bản đồ, ứng với tỷ lệ đói nghèo trên 60%. Tỷ lệ đói nghèo của các vùng này cao có thể do thực tế đây là vùng đất khô cằn nhất của Việt nam, đất cát gây ra rất nhiều khó khăn trong thâm canh nông nghiệp. Phần khá lớn dân số ở đây tham gia vào thủy sản (xem Hình 6).

Ở khu vực Tây Nguyên, bản đồ xã cho thấy tỷ lệ đói nghèo là cao nhất ở vùng phía Bắc tây Nguyên, nhất là các tỉnh Gia Lai và Kon Tum. Những xã có tỷ lệ đói nghèo thấp nhất trong 2 tỉnh này với màu xanh và vàng) là các xã dọc đường quốc lộ Bắc Nam nối từ thành phố Kon Tum với Pleiku (Quốc lộ 14) và dọc theo đường Đông Tây nối từ Pleiku ra biển (Quốc Lộ 19). Đắc Lắc có tỷ lệ đói nghèo thấp hơn Gia Lai và Kon Tum, hầu hết các xã có màu vàng hay xanh, ứng với tỷ lệ đói nghèo dưới 60%. Thành phố Buôn Ma Thuột hiện rất rõ với chấm xanh ở trung tâm của tỉnh. Đắc Lăklà vùng trồng cà phê chủ yếu của cả nước và có nhiều bước phát triển do tăng trưởng sản xuất cà phê trong thập niên (xem Hình 6 và Hình 7).

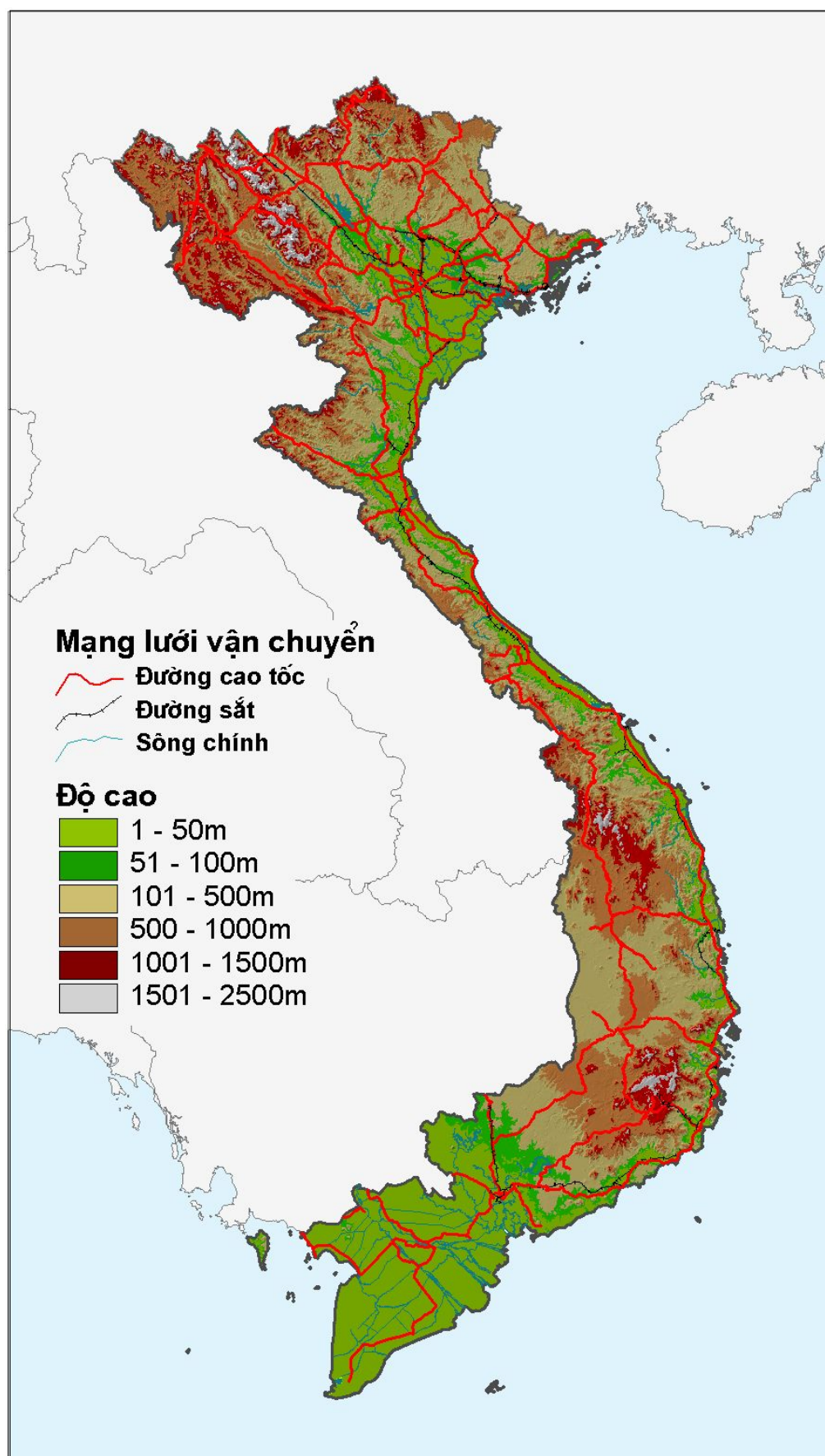
Bản đồ đói nghèo cấp xã cũng cho thấy vùng diện tích có màu đỏ ở cả ba mặt của tỉnh Lâm Đồng ở vùng Đông Nam Bộ. Đây là những vùng miền núi và rất khó tiếp cận tới đường quốc lộ hay thị trường. Đường màu xanh và vàng ở trung tâm của tỉnh theo đường quốc lộ đông bắc từ Hồ Chí Minh tới Đà Lạt (quốc lộ 20), đây là vùng xanh đậm. Sau Đà Lạt, đường quốc lộ quay về đông nam và đi ra biển, hiện rõ trên bản đồ với các xã màu vàng và màu xanh dọc theo quốc lộ. Vùng Đông Nam Bộ là vùng có diện tích màu xanh lớn nhất, cho thấy đây là vùng có tỷ lệ đói nghèo tương đối thấp (dưới 20%) ở hầu hết các tỉnh gần thành phố Hồ Chí Minh (xem Sơ đồ 6 và Sơ đồ 7). Như đề cập ở trên, sự chênh lệch rõ ràng giữa tỷ lệ đói nghèo thấp ở Đông nam Bộ (năm 1998) và tỷ lệ cao hơn ở Tây Nguyên và Nam trung Bộ phần nào do việc sử dụng các biến giả vùng trong Giai đoạn 1. Mặc dù tỷ lệ đói nghèo trung bình ở Đông Nam Bộ xấp xỉ với tỷ lệ đói nghèo bình quân của vùng tính theo VLSS, nhưng tỷ lệ đói nghèo ở những xã vùng sâu vùng xa có thể thấp hơn.

Ngược lại, ĐBSCL lại chủ yếu là màu xanh nhạt và vàng, biểu thị tỷ lệ đói nghèo từ 30-60%. Chỉ một số ít xã trong vùng có tỷ lệ đói nghèo trên 60%, tương ứng với màu cam nhạt. các xã này nằm ở tỉnh Trà Vinh và Sóc Trăng, gần cửa sông Tiền Giang và Hậu Giang các nhánh của sông Mê Kông, và ở An Giang, Kiên Giang gần biên giới Cầm Phú Chia.

Hình 6. Bản đồ đói nghèo cấp xã

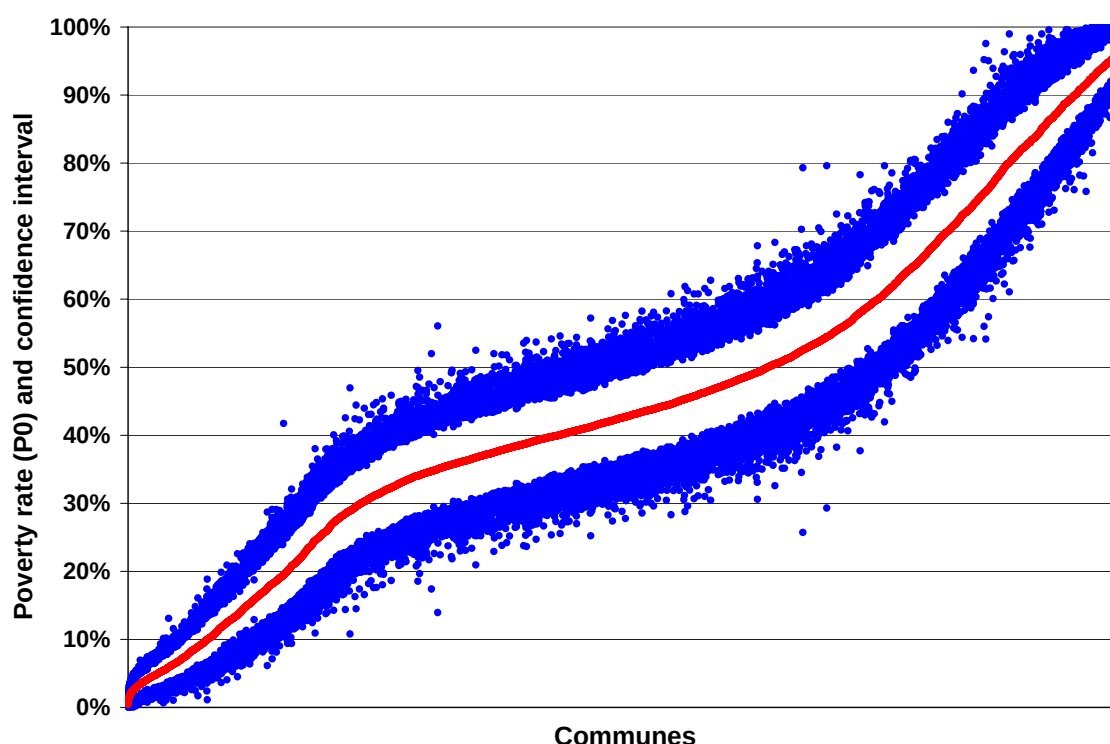


Hình 7. Bản đồ về độ cao và đường xá



Như đề cập ở trên, các ước lượng tỷ lệ đói nghèo cấp xã phải được giải thích rất thận trọng. Có một số xã có số hộ tương đối nhỏ dẫn đến sai số tương đối cao (xem Hình 8). Các khoảng tin cậy cho ước lượng đói nghèo cấp xã biến động từ dưới $\pm 1\%$ đến $\pm 27\%$, trung bình là $\pm 8.1\%$. Một nửa số xã có khoảng tin cậy nằm trong khoảng từ ± 6.6 và $\pm 10\%$. Điều này có nghĩa là $\frac{1}{4}$ số xã có khoảng tin cậy lớn hơn $\pm 10\%$. So với đói nghèo cấp huyện và tỉnh cho thấy, không có tỉnh nào có khoảng tin cậy lớn hơn $\pm 10\%$ và chỉ 15 trong tổng số 614 khoảng tin cậy của tỷ lệ đói nghèo ước lượng cho cấp huyện nằm trong khoảng này. *Rõ ràng, các ước lượng đói nghèo cấp xã cần phải thận trọng khi sử dụng, cần phải xét đến độ lớn khoảng tin cậy. Thậm chí, đối với một vài xã, chúng ta không nên sử dụng tỷ lệ đói nghèo.*

Hình 8. Tỷ lệ đói nghèo cấp xã và khoảng tin cậy

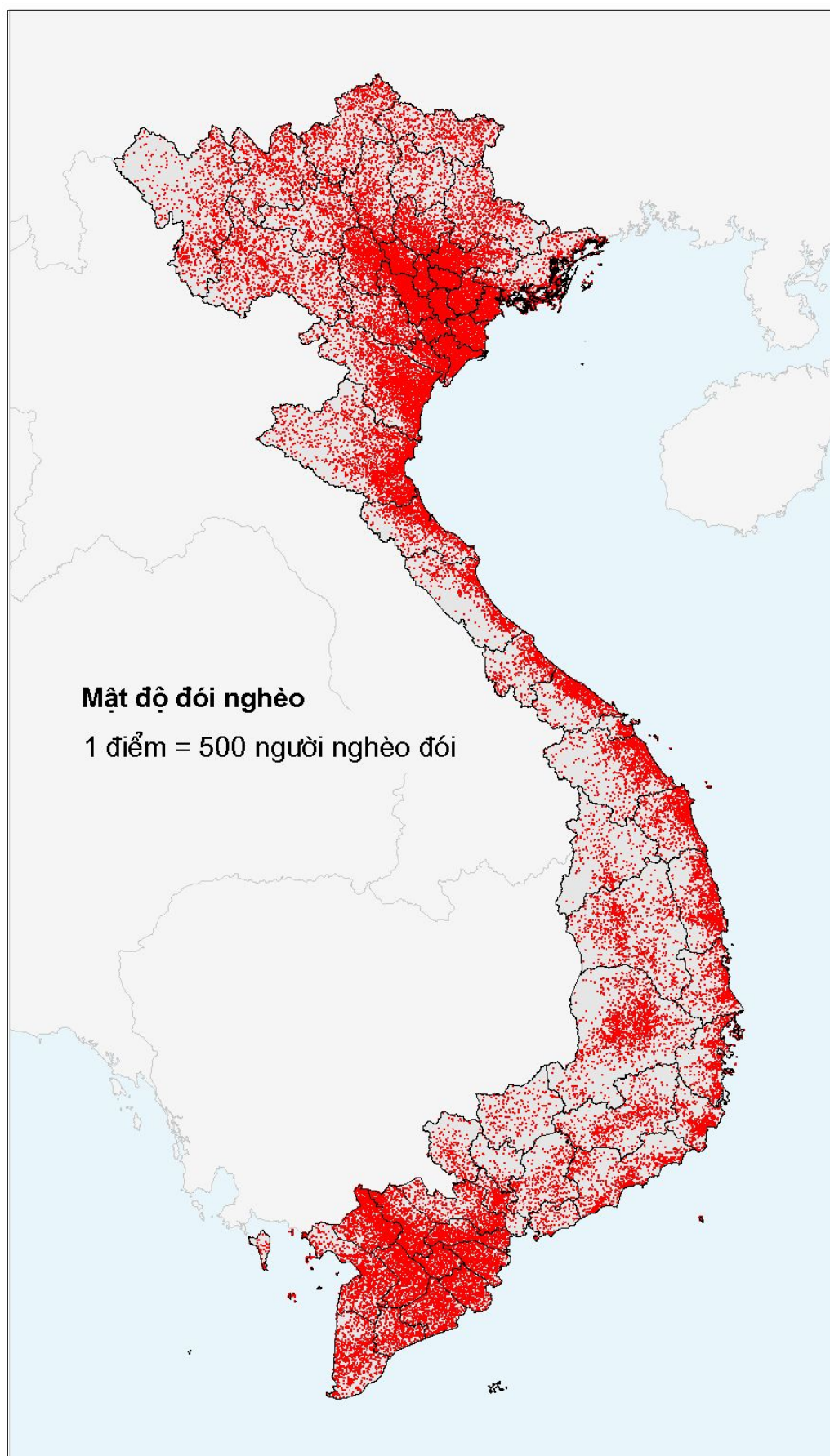


Mật độ nghèo đói

Ba bản đồ trình bày trong Hình 1, Hình 3 và Hình 6 cho ta thấy tỷ lệ đói nghèo hay tỷ lệ dân số sống dưới đường chuẩn nghèo. Một cách khác để có thể xem xét sự phân bố nghèo đói xét về mặt không gian là mật độ nghèo đói, được xác định bằng số người nghèo trong một khu vực nhất định. Bằng cách nhân tỷ lệ đói nghèo với dân số ở mỗi xã, chúng ta có thể ước tính số người nghèo ở trong xã đó, được biểu thị bằng số chấm trong các xã trên bản đồ¹¹. Hình 9 cho ta thấy mật độ đói nghèo ở Việt Nam, mỗi chấm đại diện cho 500 người nghèo.

¹¹ Chúng tôi không có thông tin sự phân bố về địa lý của các hộ trong mỗi xã, vì thế những chấm được phân bố ngẫu nhiên trong từng xã.

Hình 9. Bản đồ mật độ đói nghèo



Có điều khá ngạc nhiên là số người nghèo trên 1 km² là lớn nhất ở ĐBSH, ĐBSCL và vùng đồng bằng duyên hải. Mật độ đói nghèo thấp nhất là những vùng có tỷ lệ đói nghèo cao nhất. Điều này xảy ra bởi vì, ở những vùng có tỷ lệ đói nghèo cao thì là những vùng xa xôi có dân cư thưa thớt, và mật độ đói nghèo thấp thì tỷ lệ đói nghèo cao hơn.

Một trong ý nghĩa quan trọng của Hình 9 là nếu tất cả các chương trình xoá đói giảm nghèo được tập trung cho những vùng có tỷ lệ đói nghèo cao nhất, như Tây Bắc, Đông Bắc và Tây Nguyên, vùng đồng bằng duyên hải thì rất nhiều người nghèo sẽ không được hưởng lợi từ những chương trình này. Ý nghĩa của bản đồ này được thảo luận kỹ hơn trong Phần 6.

3.3. Các chỉ số đói nghèo về không gian khác

Phần trước đã phân tích các yếu tố không gian ảnh hưởng đến tỷ lệ đói nghèo- P0. Như đã mô tả trong Phần 2.3, có các chỉ số đói nghèo khác rất hữu ích. Độ sâu của đói nghèo (P1) hay có thể gọi là khoảng cách đói nghèo không chỉ tính số lượng người nghèo mà còn cho biết mức độ nghèo như thế nào. Trên thực tế, độ sâu của đói nghèo tương đương với tỉ lệ dân số mà mức độ nghèo được tính bằng tỉ lệ người nghèo nhân với phần trăm chênh lệch giữa đường chuẩn nghèo và mức chi tiêu bình quân đầu người của người nghèo. Độ trầm trọng của đói nghèo (P2) hay khoảng cách đói nghèo bình phương không những cho biết mức độ nghèo của các hộ như thế nào, mà còn cho biết sự phân bố thu nhập giữa những người nghèo (xem Phần 2.3 để biết thêm chi tiết).

Ở cấp độ quốc gia, giá trị ước lượng của P1 là 0.10, cho thấy người nghèo trung bình có chi tiêu thấp hơn 28% so với đường chuẩn nghèo¹²

Hình 10 là các bản đồ độ sâu của đói nghèo đến cấp tỉnh (P1) và độ trầm trọng của đói nghèo (P2) được đặt cạnh nhau để có thể dễ dàng so sánh. Rõ ràng là, về mặt không gian P1 và P2 tương đối giống nhau và cũng giống với P0 (xem Hình 3). Trong cả ba bản đồ, tỉ lệ đói nghèo cao nhất ở khu vực Tây bắc, Đông bắc, lục địa Bắc Trung bộ, phía bắc Tây Nguyên. Mức độ đói nghèo ở Đồng bằng sông Hồng và Đồng bằng sông Cửu Long là trung bình và thấp nhất ở các đô thị lớn như Hà nội và thành phố Hồ Chí Minh cũng như khu vực miền Đông Nam bộ gần thành phố Hồ Chí Minh.

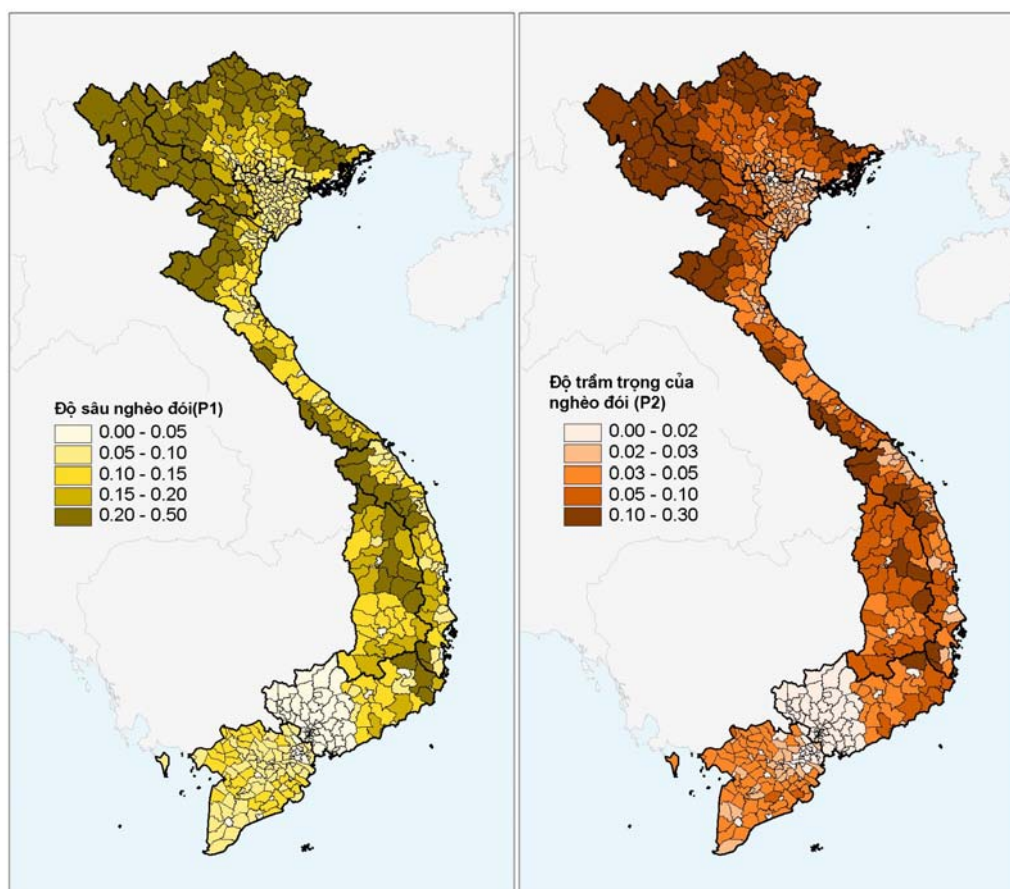
Hình 11 biểu diễn độ sâu của đói nghèo (P1) và độ trầm trọng của đói nghèo (P2) trên trục tung, tỉ lệ đói nghèo (P0) biểu diễn trên trục hoành, với mỗi điểm biểu diễn cho một tỉnh. Khi tỉ lệ đói nghèo tăng, độ sâu của đói nghèo và độ trầm trọng của đói nghèo cũng tăng theo. Mối tương quan giữa các phương pháp ước lượng đói nghèo tương đối lớn¹³. Trên thực tế, đường cong P1 tăng lên khi P0 tăng cho thấy, khi tỉ lệ đói nghèo tăng, tỉ lệ khoảng cách giữa

¹² $P1=P0 \cdot G$, trong đó G là chênh lệch giữa đường chuẩn nghèo và chi tiêu bình quân đầu người trung bình của người nghèo, được biểu thị như là tỷ lệ của đường chuẩn nghèo.

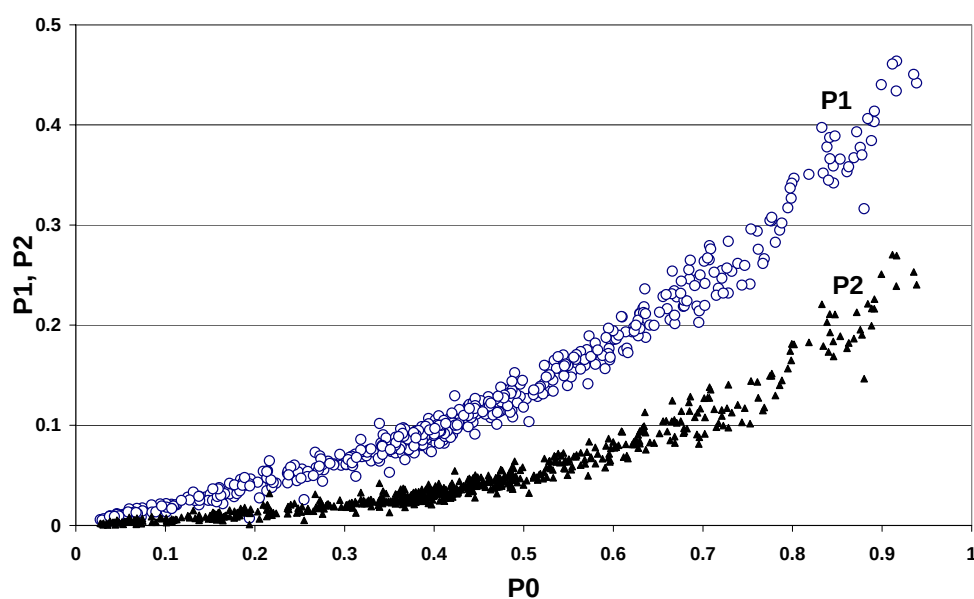
¹³ đường xu thế bậc hai dựa trên P0 có chỉ số R^2 bằng 0.98 đối với P₁ và bằng 0.96 đối với P₂

đường đói nghèo và mức chi tiêu bình quân đầu người của các hộ nghèo cũng tăng lên¹⁴.

Hình 10. Bản đồ độ sâu của đói nghèo (P_1) và độ trầm trọng của đói nghèo (P_2) của mỗi huyện



Hình 11. Độ sâu của đói nghèo (P_1) và độ trầm trọng của đói nghèo (P_2) biểu diễn như một hàm tỉ lệ đói nghèo (P_0) ở mỗi huyện



¹⁴ Xem chú thích 11. Nếu khoảng cách đói nghèo trung bình là hằng số, P_1 sẽ có tương quan tuyến tính cùng chiều với P_0 . Thực tế là đường cong đi lên cho biết khoảng cách đói nghèo bình quân cũng phải tăng khi chúng ta dịch chuyển tới huyện nghèo hơn.

3.4. Bất bình đẳng giữa các vùng

Như đã phân tích trong phần 2.4, phương pháp ước lượng diện tích nhỏ thường được sử dụng để ước lượng tỉ lệ đói nghèo (lập bản đồ đói nghèo), song phương pháp này cũng có thể được sử dụng để tính toán các chỉ số bất bình đẳng. Trong khi các chỉ số đói nghèo tập trung vào những người sống dưới mức chuẩn nghèo, các chỉ số bất bình đẳng xem xét sự phân bố của cả người nghèo và người không nghèo. Trong phân tích này, chúng tôi tập trung vào ba chỉ số đo lường chính : hệ số Gini, chỉ số bất bình đẳng Theil L và chỉ số bất bình đẳng Theil T. Hai chỉ số Theil thuộc tập hợp các chỉ số Entropy tổng hợp (Generalized Entropy), đôi khi được viết là GE(0) và GE(1).

Hệ số Gini

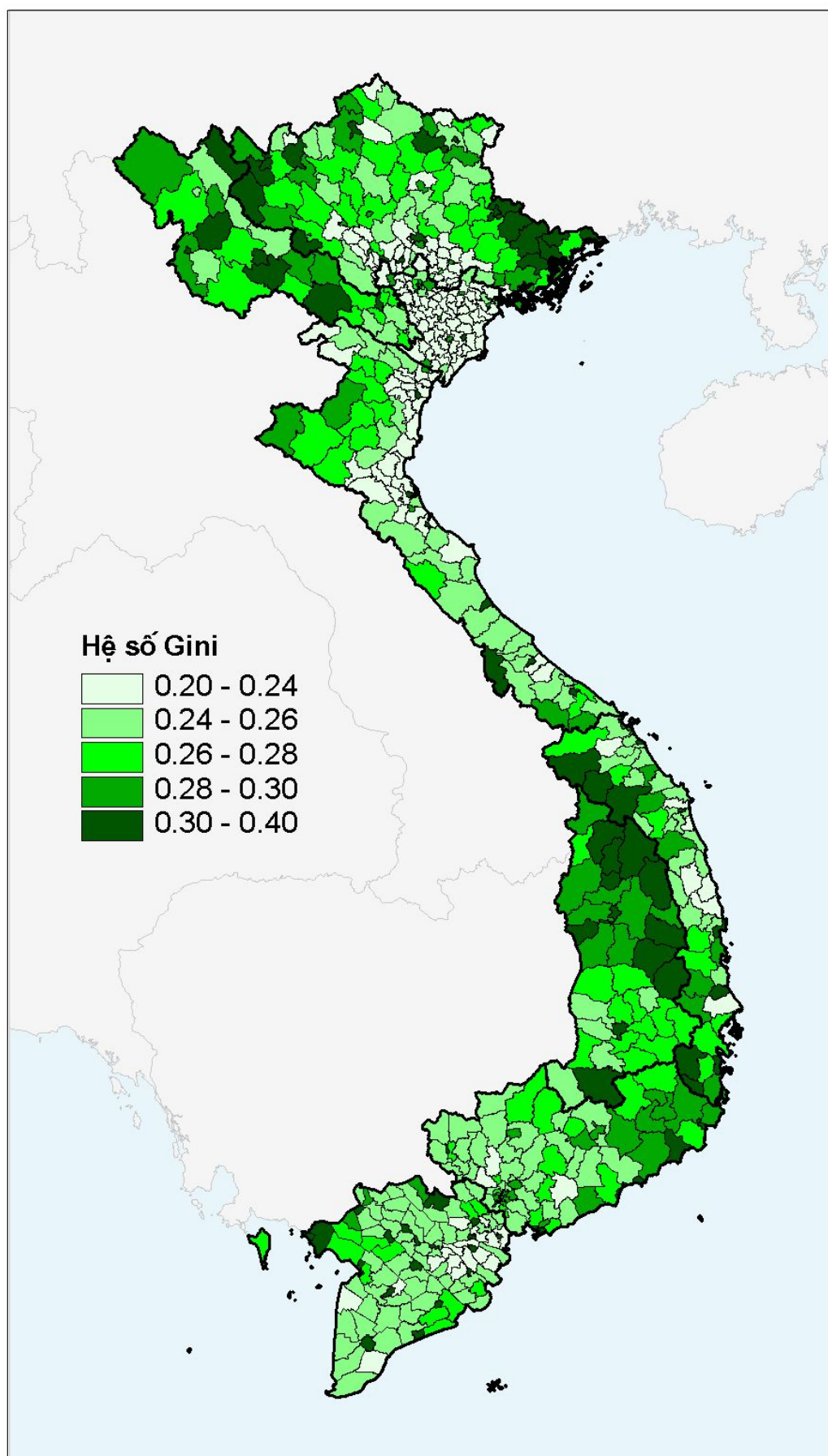
Hệ số Gini là một đơn vị đo lường sự bất bình đẳng, dao động giữa 0 (khi tất cả mọi người có cùng một mức chi tiêu hoặc thu nhập) và 1 (khi một người tất cả!). Do đó, hệ số Gini cao tức là sự bất bình đẳng càng cao. Đối với phần lớn các nước đang phát triển, hệ số Gini dao động từ 0,3 – 0,6. Theo phân tích của chúng tôi, hệ số Gini của cả nước là 0,323. cho thấy mức bất bình đẳng về chi tiêu bình quân đầu người tương đối thấp.

Giống như các đơn vị đo lường sự bất bình đẳng khác, hệ số Gini thường nhỏ hơn đối với các diện tích nhỏ, như tỉnh hoặc huyện hơn trên phạm vi quốc gia. Nguyên nhân là do các hộ gia đình ở các vùng nhỏ thường có mức sống giống nhau hơn so với các hộ trên phạm vi toàn quốc.

Hình 12 cho thấy mức độ bất bình đẳng trong mức chi tiêu bình quân đầu người tính theo hệ số Gini đến cấp tỉnh. Khu vực có tỉ lệ bất bình đẳng thấp (đoạn có màu trắng) gồm có Đồng bằng sông Hồng, một số vùng đồng bằng phía Đông bắc, một số huyện ven biển thuộc Duyên hải Bắc Trung bộ, một số huyện ở Đồng bằng sông Cửu Long và một vài huyện thuộc Duyên hải Nam Trung bộ. Mức độ bất bình đẳng về chi tiêu lớn nhất ở các khu vực đô thị lớn, đặc biệt ở Hà nội và thành phố Hồ Chí Minh và khu vực miền núi bao gồm cả Đông bắc, Tây bắc và Tây Nguyên.

Không ngạc nhiên khi các khu vực đô thị lớn lại có tỉ lệ bất bình đẳng cao, vì ở các vùng này thường có các hộ giàu nhất cả nước và cả những hộ mới nhập cư thu nhập của họ không cao hơn thu nhập của hộ nông thôn. Tỉ lệ bất bình đẳng thấp ở khu vực Đồng bằng sông Hồng và các vùng duyên hải cũng không bất ngờ. Điều này có được là nhờ nền nông nghiệp thâm canh và tỉ lệ dân số sống phụ thuộc vào nông nghiệp chiếm phần lớn. Triển vọng làm nông nghiệp trên các mảnh ruộng có hệ thống thủy lợi tương đối đồng nhất, việc phân bổ đất hợp tác xã giữa các hộ gia đình được tiến hành với mục đích duy trì sự bình đẳng giữa các hộ.

Hình 12. Bản đồ bất bình đẳng tính theo hệ số Gini



Đồng bằng sông Cửu Long cũng đặc trưng bởi nền nông nghiệp thâm canh cao và tỉ lệ dân số sống bằng nông nghiệp chiếm phần lớn, song lại có sự biến động lớn về quy mô đất canh tác cũng như sự góp mặt của một số hộ không có đất canh tác mà chỉ chủ yếu trông chờ vào bán sức lao động nông nghiệp.

Có lẽ điều bất ngờ nhất là sự bất bình đẳng cao tại một số vùng thuộc khu vực Đông bắc, Tây Bắc và Tây Nguyên. Một lời giải thích khá hợp lý là các khu vực này thường có nhiều nông dân rất nghèo, những người này hầu hết là người dân tộc thiểu số và có một số hộ giàu có thu nhập từ hoạt động buôn bán, sản xuất nông nghiệp có tính thương mại cao (như chăn nuôi gia súc), hoặc là lao động được trả lương, bao gồm cả cán bộ nhà nước. Tuy nhiên, Đắk Lắk là một ngoại lệ. Chúng tôi dự đoán có sự khác biệt lớn giữa nông dân trồng cà phê giàu có và những nông dân nghèo khác, song Đắk Lắk là một trong một vài tỉnh miền núi có tỉ lệ bất bình đẳng tương đối thấp.

Chỉ số Theil L và Theil T

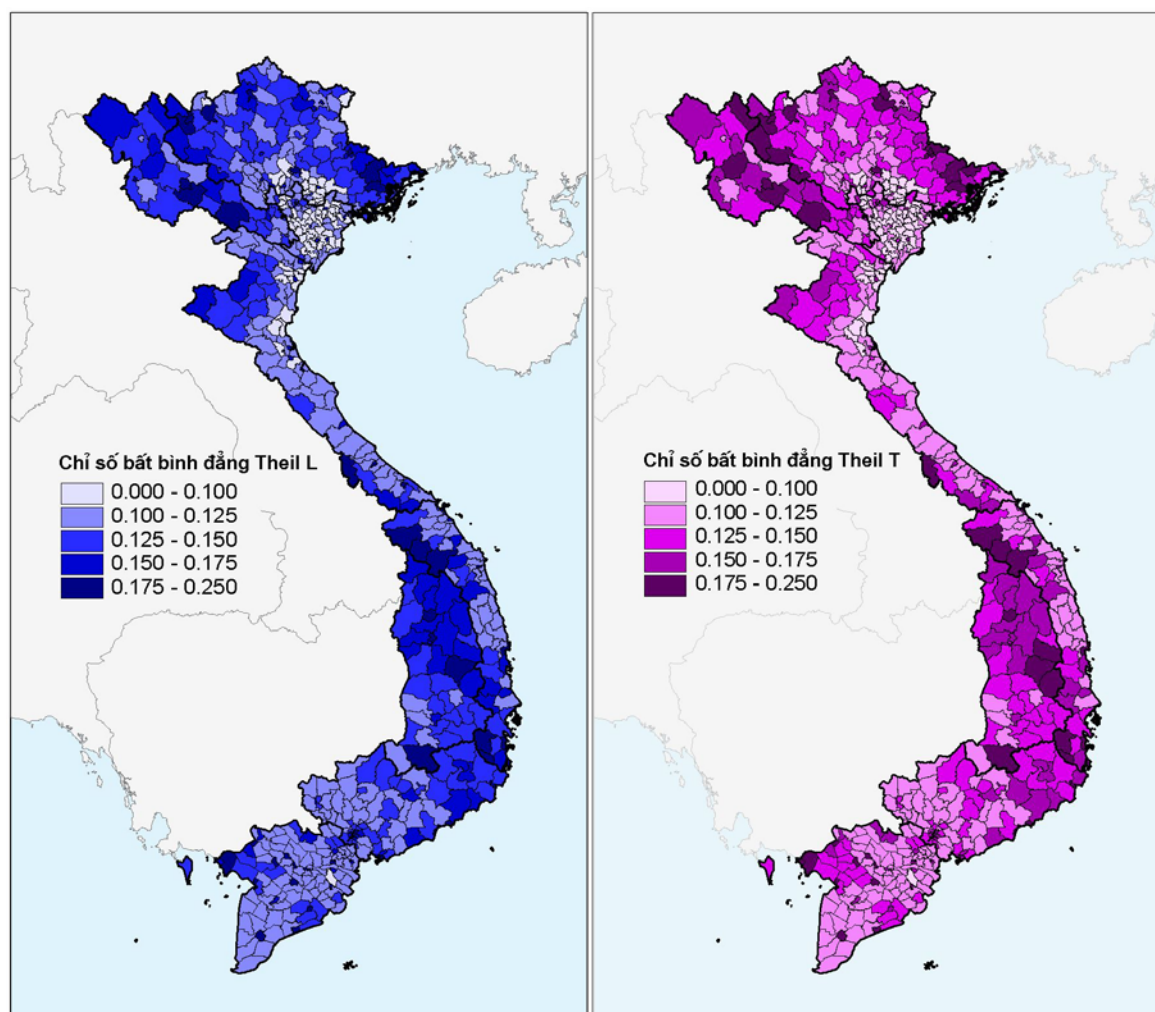
Chỉ số Theil L dao động từ 0 (bình đẳng tuyệt đối) và ∞ (bất bình đẳng tuyệt đối), mặc dù rất hiếm khi nó vượt quá 1. Giống như hệ số Gini, chỉ số Theil cao có nghĩa là sự bất bình đẳng trong phân bổ chi tiêu càng lớn (hoặc thu nhập). Tuy nhiên, chỉ số Theil L cho thấy rõ hơn sự phân bổ chi tiêu giữa người nghèo hơn so với hệ số Gini.

Chỉ số Theil T dao động từ 0 và $\log(N)$, trong đó N là dân số. Chỉ số Theil T cho biết quyền số cân bằng đối với tất cả các phần của sự phân bổ. Công thức tính hai chỉ số Theil trong Phần 2.4.

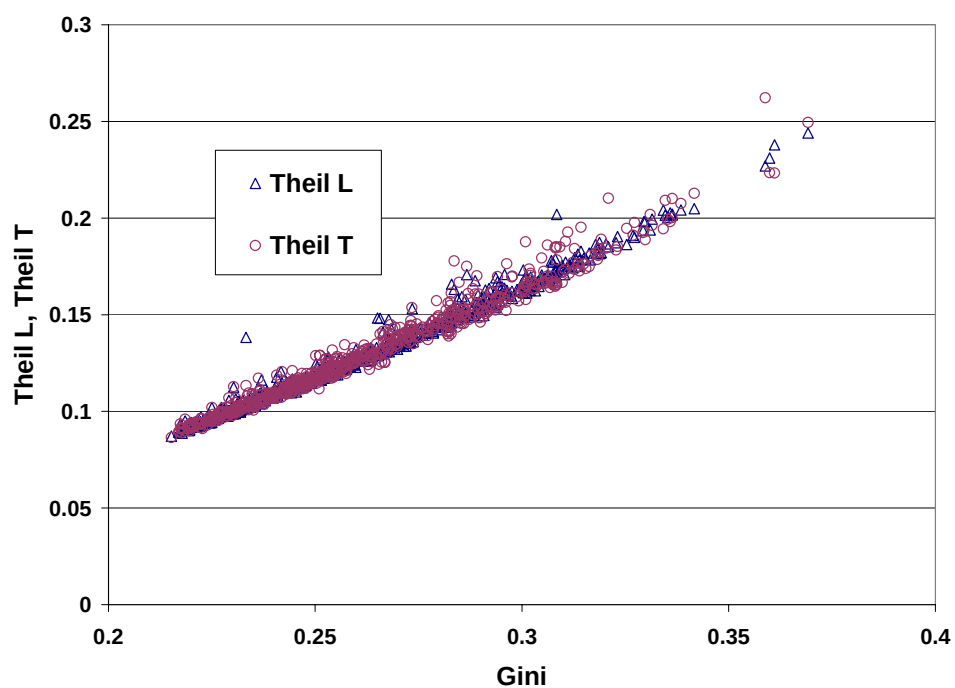
Hình 13 là bản đồ bất bình đẳng cấp huyện, miêu tả chỉ số Theil L, Hình 14 là bản đồ bất bình đẳng cấp huyện, biểu thị qua chỉ số Theil T. Mặc dù công thức tính khác nhau, bản đồ biểu thị sự bất bình đẳng sử dụng hai chỉ số Theil đều cho kết quả như nhau và giống với bản đồ về hệ số Gini. Trong cả ba bản đồ, sự bất bình đẳng thấp nhất ở Đồng bằng sông Hồng và một số vùng duyên hải miền Nam Trung bộ, bất bình đẳng ở mức độ trung bình ở khu vực Đồng bằng sông Cửu Long và lớn nhất ở các khu đô thị lớn, khu vực miền núi phía Bắc và Tây Nguyên.

Hình 15 và 16 biểu diễn mối liên hệ giữa ba đơn vị đo lường tỉ lệ bất bình đẳng, trong đó mỗi dấu chấm đại diện một huyện. Một mặt, các sơ đồ này cho thấy có một mối liên hệ tuyến tính giữa hệ số Gini và mặt khác là tương quan hai chỉ số Theil và mối tương quan này khá chặt chẽ. Đối với mối liên hệ giữa hệ số Gini và chỉ số Theil L, $R^2 = 0,98$, trong khi trong mối liên hệ giữa hệ số Gini và chỉ số Theil T, $R^2 = 0,97$. Điều này giúp giải thích lý do bản đồ bất bình đẳng của ba hệ số này rất giống nhau.

Hình 13. Bản đồ bất bình đẳng qua chỉ số Theil L và Theil T



Hình 14. Chỉ số bất bình đẳng Theil L và Theil T là hàm của hệ số Gini cho mỗi huyện



Phân tích sự bất bình đẳng

Có phải bất bình đẳng chủ yếu là do sự khác biệt giữa các tỉnh hay sự chênh lệch giữa các hộ trong mỗi tỉnh ? Không giống hệ số Gini, chỉ số Theil L và Theil T của sự bất bình đẳng có thể được phân tách một cách chính xác thành các “nhóm nhỏ hơn”. Ví dụ, chỉ số Theil cho cả nước bằng mức trung bình (có quyền số) của các chỉ số cấp tỉnh (yếu tố “trong tỉnh”) cộng với chỉ số Theil của sự bất bình đẳng trong chỉ tiêu bình quân cấp tỉnh (yếu tố “giữa tỉnh”). Yếu tố “giữa các tỉnh” cho thấy bất bình đẳng xảy ra khi tất cả người dân trong một tỉnh có cùng mức chi tiêu bằng mức trung bình của tỉnh, trong khi yếu tố “trong tỉnh” có tính đến sự bất bình đẳng trong phạm vi của tỉnh, song không tính đến sự bất bình đẳng của các trung bình tỉnh.

Bảng 7 phân tích chỉ số Theil L và Theil T sử dụng các tỉnh làm nhóm nhỏ. Yếu tố gây bất bình đẳng giữa tỉnh chiếm ít hơn 1/4 bất bình đẳng cấp quốc gia (13% đối với chỉ số Theil L và 24% đối với chỉ số Theil T). Một lượng khác lớn hơn 3/4 là do yếu tố bất bình đẳng trong mỗi tỉnh. Tầm quan trọng của các kết quả phân tách này giống với kết quả của Kanbur (2002) đối với các nước đang phát triển, yếu tố giữa tỉnh chiếm khoảng 15% trong tổng tỉ lệ bất bình đẳng cả nước¹⁵.

Bảng 7. Phân tích tỷ lệ bất bình đẳng thành các nhóm giữa tỉnh và trong cùng tỉnh

Chỉ số bất bình	Biến	Tỉ lệ bất bình	Tỉ lệ giữa tỉnh	Tỉ lệ trong tỉnh
Chỉ số Theil L	Giá trị chỉ số	0,192	0,045	0,147
	Tỉ lệ trong tổng số	100%	13%	87%
Chỉ số Theil T	Giá trị chỉ số	0,206	0,050	0,155
	Tỉ lệ trong tổng số	100%	24%	76%

Bảng 8 phân tách chỉ số Theil L và T sử dụng huyện là “nhóm nhỏ”. Yếu tố gây bất bình đẳng giữa huyện giảm xuống dưới 1/3 (34% đối với chỉ số Theil L và 36% đối với chỉ số Theil T). 2/3 tỉ lệ bất bình đẳng quốc gia liên quan tới yếu tố bất bình đẳng trong mỗi tỉnh. Chúng tôi hy vọng yếu tố bất bình đẳng giữa các huyện tăng lên khi phân vùng địa lý càng nhỏ. Điều này cho thấy rằng, ngược lại với quan điểm hiện tại của Việt Nam, không phải các tỉnh và các huyện phát triển nhanh sẽ tạo ra sự chênh lệch lớn về mức sống.

¹⁵ Rõ ràng, độ lớn tương đối giữa các thành phần hay trong các thành phần phụ thuộc vào số nhóm (tỉnh, huyện hay các đơn vị hành chính khác) liên quan. Số nhóm càng lớn, bất bình đẳng giữa các thành phần càng lớn

Bảng 8. Phân tách sự bất bình đẳng thành các nhóm giữa huyện và trong huyện

Chỉ số bất	Biến	Tỉ lệ bất bình đẳng	Tỉ lệ giữa huyện	Tỉ lệ trong huyện
Chỉ số Theil L	Giá trị chỉ số	0.193	0.067	0.127
	Tỉ lệ trong tổng số	0.204	0.073	0.131
Chỉ số Theil T	Giá trị chỉ số			
	Tỉ lệ trong tổng số			

3.5. Mối liên hệ giữa thu nhập, đói nghèo và bất bình đẳng

Phần 3.2 và 3.4 đã nghiên cứu các vấn đề về đói nghèo và bất bình đẳng theo không gian. Trong phần này, chúng tôi nghiên cứu mối liên hệ giữa đói nghèo, bất bình đẳng, sự đô thị hoá và mức chi tiêu bình quân đầu người trung bình ở cấp huyện. Để giảm số lượng biến và do mối tương quan chặt chẽ giữa các đại lượng ước lượng đói nghèo, chúng tôi sẽ sử dụng P_0 để biểu thị tỉ lệ đói nghèo. Tương tự như vậy, do tất cả ba đại lượng ước lượng bất bình đẳng cũng có mối tương quan chặt chẽ, chúng tôi sẽ sử dụng hệ số Gini để biểu diễn tỉ lệ bất bình đẳng.

Trong Hình 18, chúng tôi vẽ biểu đồ tỉ lệ đói nghèo là một hàm số của mức chi tiêu bình quân đầu người cấp huyện, trong đó, mỗi dấu chấm là một huyện. Chúng tôi hy vọng khi chi tiêu bình quân đầu người tăng, tỉ lệ đói nghèo sẽ giảm. Tuy nhiên, điều ngạc nhiên là tỉ lệ đói nghèo phụ thuộc rất nhiều vào mức chi tiêu bình quân đầu người của huyện đó. Đặc biệt đối với các huyện nghèo, mối liên hệ giữa hai yếu tố này rất chặt chẽ. Đường xu hướng phương trình bậc hai giải thích 96% sự biến động của đói nghèo. Điều này cho thấy tỉ lệ đói nghèo trong một huyện là một hàm số của mức chi tiêu bình quân đầu người tại huyện đó và tỉ lệ bất bình đẳng trong một huyện chỉ đóng vai trò nhỏ tác động đến tỉ lệ đói nghèo.

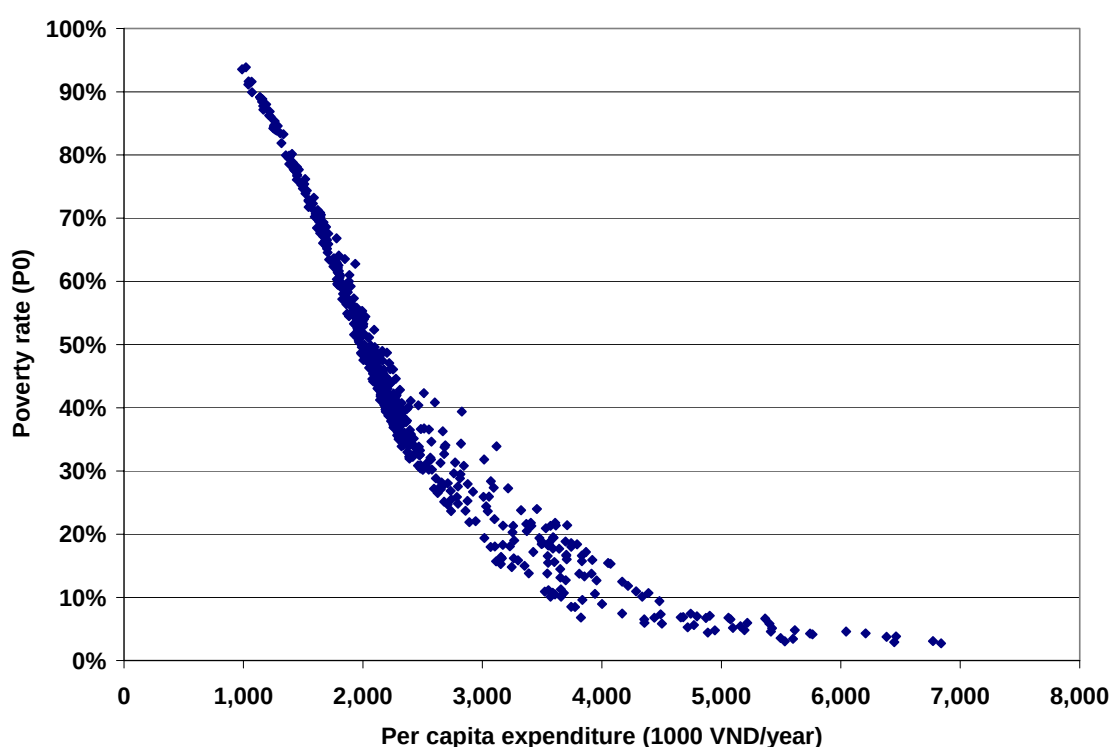
Hình 19 biểu diễn mối liên hệ giữa hệ số Gini và mức chi tiêu bình quân đầu người của huyện. Ở Việt Nam và một số nước khác trên thế giới, người ta tin rằng nếu thu nhập tăng, khoảng cách giữa người nghèo và người giàu sẽ càng lớn. Số liệu ở đây khẳng định quan điểm ở một số cấp độ. Đường tuyến tính trên đồ thị cho biết khi hệ số Gini tăng từ 0,25 – 0,30 thì chi tiêu bình quân đầu người tăng từ 1 triệu đồng/năm đến 7 triệu đồng/năm. Điều này có thể thấy trong các nghiên cứu trên thế giới, ở mức thu nhập thấp, thì nếu thu nhập cao kéo theo bất bình đẳng cao, song ở một vài điểm, thu nhập càng cao thì xu hướng bất bình đẳng lại giảm. Đường U ngược này gọi là đường cong Kuznets. Vì Việt Nam là nước có thu nhập bình quân thấp nên đường cong Kuznets sẽ cho thấy tương quan cùng chiều giữa thu nhập và bất bình đẳng qua các năm và giữa các huyện.

Tuy nhiên, mối liên hệ giữa bất bình đẳng và chi tiêu bình quân đầu người trong Hình 16 không phải là mối quan hệ chùng chiều đơn giản. Nhiều huyện thu nhập thấp cũng có tỉ lệ bất bình đẳng cao. Trên thực tế, các vùng có tỉ lệ bất bình đẳng cao thường là các huyện tương đối nghèo với mức chi tiêu bình quân đầu người dưới 4 triệu đồng/năm. Hơn nữa, các

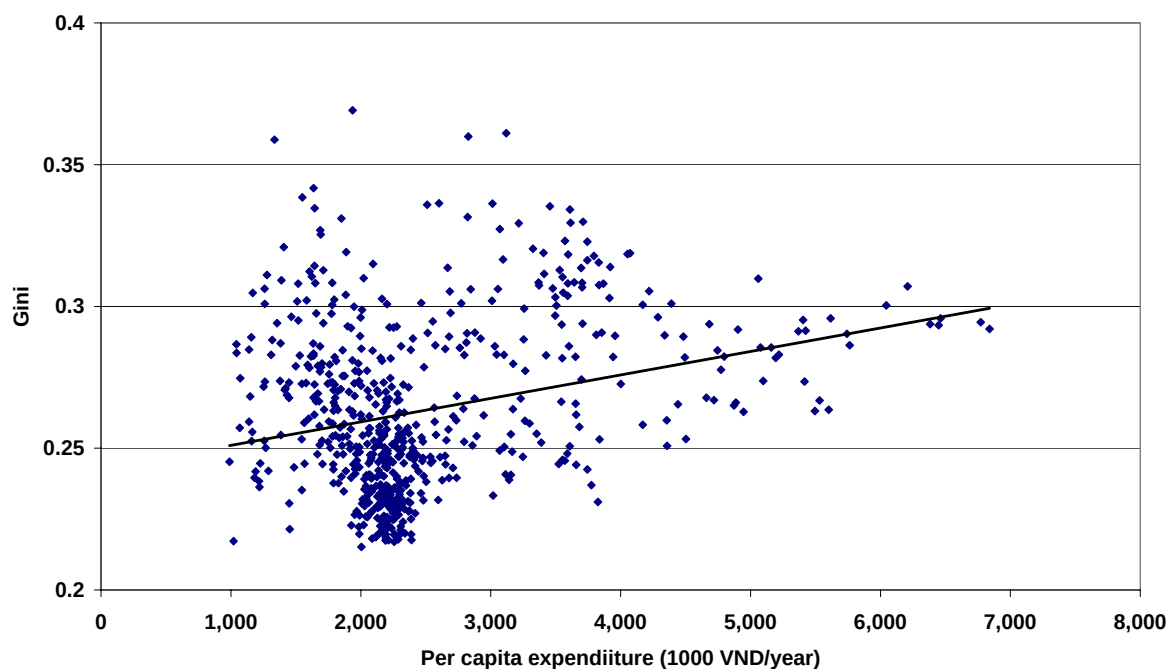
huyện thu nhập thấp có độ dao động lớn về bất bình đẳng, trong khi các huyện thu nhập cao có thì hệ số Gini nằm ở mức xấp xỉ 0,3.

Mối liên hệ giữa tỉ lệ đói nghèo (P_0) và bất bình đẳng (hệ số Gini) được biểu diễn trong Hình 20. Ta thấy đường biểu diễn hình chữ U trong đó tỉ lệ bất bình đẳng cao nhất là ở các huyện nghèo nhất và các huyện giàu nhất. Điều này có thể liên quan tới xu hướng ở hình 12 trong đó, các vùng có tỉ lệ bất bình đẳng cao ở khu vực miền núi (nơi có tỉ lệ đói nghèo cao) và các trung tâm đô thị lớn (nơi có tỉ lệ nghèo thấp). Mặc dù xu hướng này củng cố một mối liên hệ phương trình bậc hai (cong), nhưng sự tương quan này vẫn tương đối yếu ($R^2 = 0,12$).

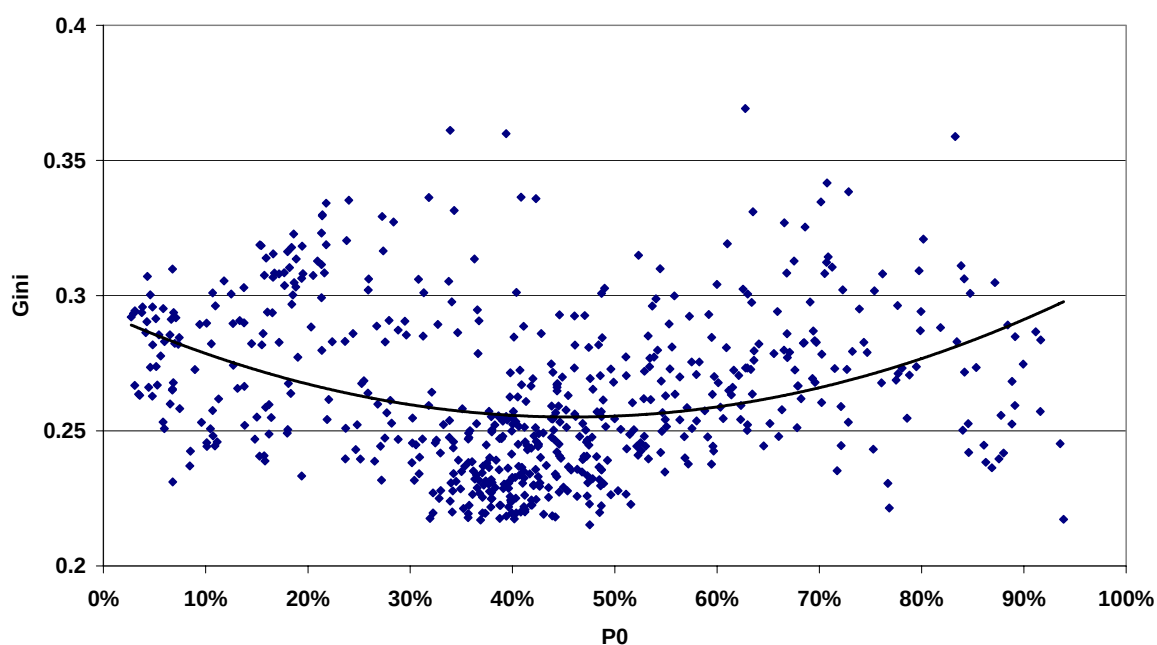
Hình 18. Tỉ lệ đói nghèo (P_0) như một hàm số của chi tiêu bình quân đầu người



Hình 19. Hệ số Gini bất bình đẳng như một hàm số của chi tiêu bình quân đầu người



Hình 20. Hệ số Gini bất bình đẳng như một hàm số của tỉ lệ đói nghèo (P_0)

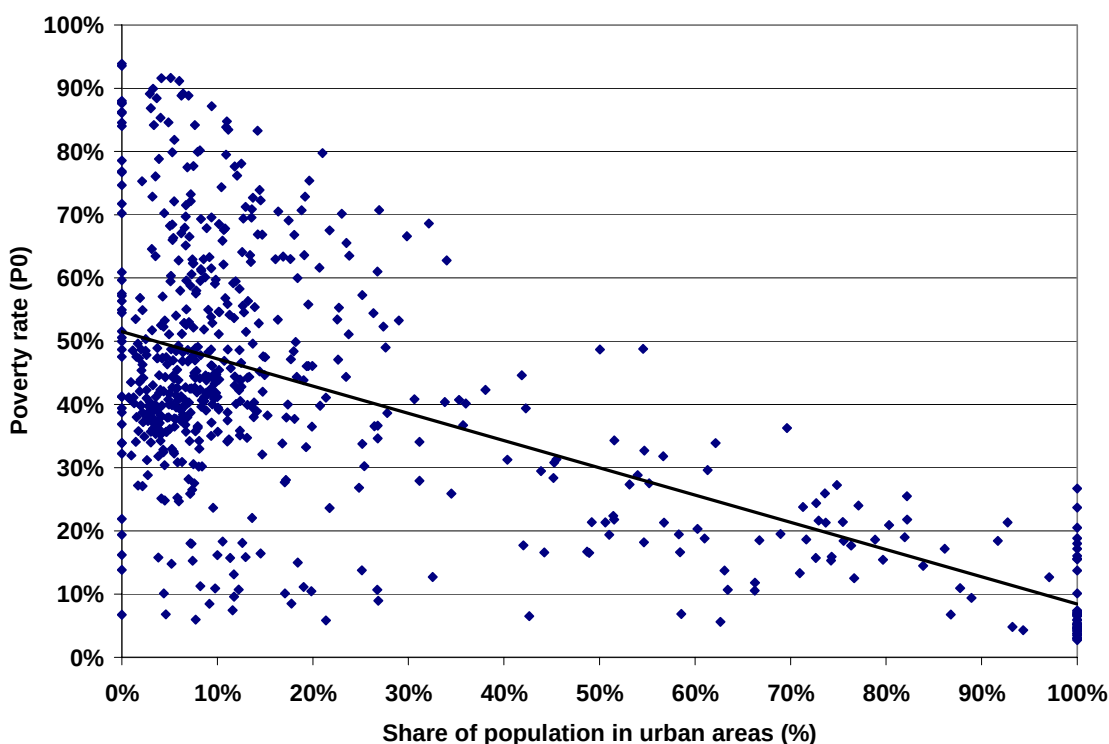


Tỉ lệ đói nghèo ở nông thôn cao hơn thành thị ở hầu hết các nước trên thế giới đã tiến hành nghiên cứu. Điều này đã được khẳng định ở Việt Nam bằng nhiều cuộc điều tra khác nhau (có thể xem GSO, 2000). Tuy nhiên, với phương pháp ước lượng quy mô nhỏ, chúng tôi có thể kiểm tra tỉ lệ đói nghèo ở nhiều vùng nông thôn và thành thị để cung cấp một bức tranh

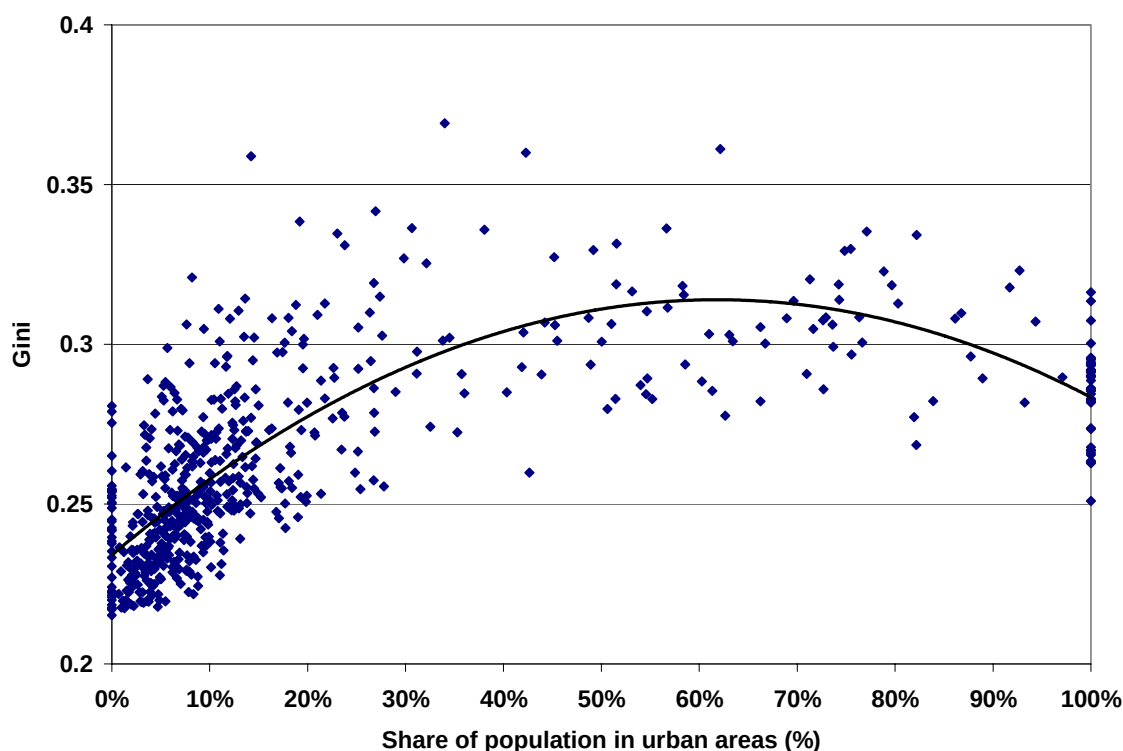
chi tiết hơn về mối liên hệ giữa đô thị hoá và đói nghèo. Hình 21 biểu diễn mối liên hệ giữa dân số sống ở khu vực thành thị và tỉ lệ đói nghèo (P0) của các huyện theo đường tuyến tính. Sơ đồ này cho thấy mối tương quan ngược: phần lớn các vùng chủ yếu là nông thôn có tỉ lệ đói nghèo từ 30 – 60%, trong khi phần lớn các vùng chủ yếu là thành thị có tỉ lệ đói nghèo dưới 30%. Đồng thời, ta cũng thấy sự khác biệt lớn về tỉ lệ đói nghèo giữa các vùng nông thôn. Một số huyện có phần lớn dân số nông thôn nhưng tỷ lệ đói nghèo tương đương với các huyện thành thị. Điều này chứng tỏ rằng trong một số trường hợp, đói nghèo có thể giảm đáng kể ở các vùng nông thôn. Dựa vào các bản đồ trình bày ở trên, ta thấy nhiều vùng nông thôn thu nhập cao là ở khu vực miền đông Nam bộ do, lợi ích từ tiếp cận với thị trường lao động và hàng hoá ở thành phố Hồ Chí Minh.

Mối liên hệ giữa đô thị hoá và bất bình đẳng khác nhau hoàn toàn. Trong Hình 22, tỉ lệ bất bình đẳng (theo tính toán với hệ số Gini) tương đối thấp ở các vùng phần lớn dân số sống ở nông thôn và cũng thấp đối với những huyện thành thị. Những huyện có tỉ lệ bất bình đẳng cao là những vùng có cả dân số nông thôn và thành thị, với mức độ đô thị hoá dao động từ 20 – 80%. Các kết quả này khẳng định quan điểm chung là các huyện thành thị có tỉ lệ bất bình đẳng cao hơn các huyện nông thôn, song lại chứng tỏ rằng xu hướng này khá phức tạp ở những huyện có cả dân số nông thôn và thành thị và có tỉ lệ bất bình đẳng cao. Đường phương trình bậc hai này cho thấy sự bất bình đẳng cao nhất khi mức độ đô thị hoá đạt 60%.

Hình 21. Tỷ lệ đói nghèo (P0) là hàm của mức độ đô thị hoá



Hình 22. Hệ số bất bình đẳng Gini là hàm của đô thị hoá



3.6 So sánh với ước lượng về đói nghèo của MOLISA

Trong phần này, chúng tôi so sánh các ước lượng tỷ lệ đói nghèo (P0) từ nghiên cứu áp dụng phương pháp ước lượng khu vực nhỏ với tỷ lệ đói nghèo ước lượng của Bộ Lao động và Thương binh Xã hội (MOLISA). Như miêu tả ở trên, có một số cách định nghĩa về đói nghèo và phương pháp thu thập số liệu khác nhau. Sự khác nhau đó có thể tổng hợp như sau:

Định nghĩa về đói nghèo của chúng tôi sử dụng như là một chỉ số về của cải (welfare indicator): chi tiêu bình quân đầu người, gồm giá trị lương thực thực phẩm ở mức đủ sống và các giá trị nhà ở. Ngược lại, MOLISA sử dụng thu nhập bình quân đầu người như là chỉ số về của cải.

Để có thể điều chỉnh sự khác nhau giữa các vùng trong mức chi tiêu, chúng tôi sử dụng một tập chỉ số giá vùng theo tháng của Tổng Cục Thống kê cho phân tích VLSS 1998. Những chỉ số giá này dựa trên chi phí của rổ hàng hoá cơ bản ở nông thôn và thành thị của mỗi vùng. Ngược lại, MOLISA điều chỉnh chi tiêu của vùng bằng cách biểu thị thu nhập bình quân đầu người theo số túi gạo sẽ mua với mức giá ở vùng đó¹⁶.

¹⁶ Điều này Tương đương với chỉ có một chỉ số giá mà chỉ gồm 1 hàng hoá là gạo trong rổ hàng hoá tiêu thụ.

Đường chuẩn nghèo chúng tôi sử dụng là “đường đói nghèo chung”, ở mức chi tiêu thực tế là 1,789 triệu đồng/người/năm. MOLISA định nghĩa đường chuẩn nghèo theo số lượng gạo, mặc dù có sự khác nhau giữa các tỉnh.

Chúng tôi định nghĩa tỷ lệ đói nghèo là phần trăm dân số trong hộ có chi tiêu bình quân đầu người dưới đường chuẩn nghèo. MOLISA định nghĩa tỷ lệ đói nghèo là phần trăm số hộ dưới đường chuẩn nghèo.

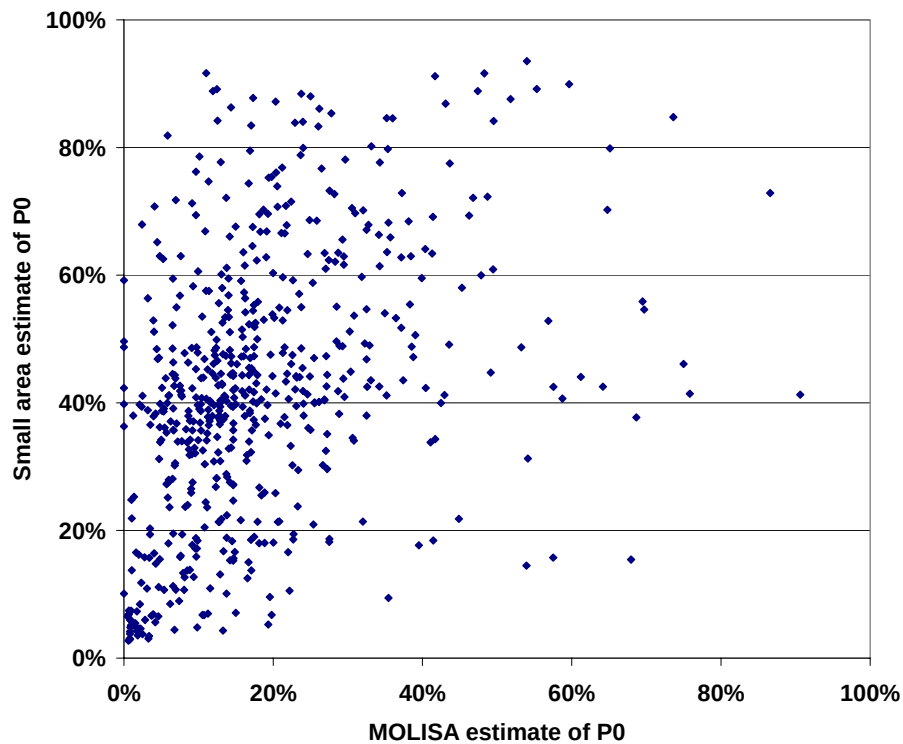
Ước lượng đói nghèo của mỗi huyện dựa trên các đặc điểm của hộ trong huyện đó theo Tổng điều tra nhà ở và dân số năm 1999, với tương quan giữa chi tiêu bình quân đầu người và những đặc điểm của những hộ này trong VLSS 1998. Ước lượng của MOLISA dựa trên đánh giá của các cán bộ thực địa của MOLISA trong mỗi xã, áp dụng nguyên tắc chỉ đạo của tỉnh và toàn quốc về việc xác định hộ nghèo (miêu tả công tác thực địa, xem Conway, 2001).

Liệu các phương pháp khác nhau có cho các kết quả ước lượng về tỷ lệ đói nghèo (P_0) ở cấp huyện khác nhau hay không? Như trình bày trong Sơ đồ 23, tỷ lệ đói nghèo của MOLISA nhà chung là thấp hơn tỷ lệ đói nghèo ước lượng từ phương pháp ước lượng theo diện tích nhỏ (small-area estimation method) dùng cho báo cáo này. Giá trị trung vị của các ước lượng tỷ lệ đói nghèo của MOLISA là 15%, so với 41% theo ước tính của chúng tôi. Sự khác nhau này không có gì ngạc nhiên vì chúng tôi dựa trên hai đường chuẩn nghèo khác nhau. Với rất nhiều các quan điểm về làm thế nào để xây dựng đường chuẩn nghèo, các cuộc tranh luận về “tỷ lệ đói nghèo” chính xác cũng có ý nghĩa. Tuy nhiên việc xem xét đói nghèo về mặt không gian của hai phương pháp có thống nhất là điều quan trọng hơn.

Điều đáng ngạc nhiên là có sự tương quan rất yếu giữa các ước lượng tỷ lệ đói nghèo của MOLISA và các tỷ lệ đói nghèo ước tính trong nghiên cứu này (R^2 của đường ước lượng tuyến tính -linear trendline- chỉ bằng 0.17). Để minh họa sự không thống nhất trong các ước lượng, chúng tôi đề cập tới hai huyện mà mâu thuẫn giữa hai phương pháp là lớn nhất. Trong góc bên trái phía trên của Sơ đồ 23 là chấm biểu thị cho huyện Bát Xát ở Lào Cai. Theo ước lượng của chúng tôi, tỷ lệ đói nghèo của Bát Xát là khoảng 82%. Ngược lại, tỷ lệ đói nghèo theo ước tính của MOLISA cho huyện này là dưới 6%. Ở góc phía dưới bên phải là huyện Nha Trang, tỉnh Khánh Hoà. Ước lượng tỷ lệ đói nghèo của MOLISA cho huyện Nha Trang là 68 %, trong khi đó, tỷ lệ trong báo cáo này chỉ là 15%.

Rõ ràng, việc lựa chọn các ước lượng đói nghèo có thể tạo ra sự khác biệt lớn trong mục tiêu của các chương trình xoá đói giảm nghèo. Cần có nghiên cứu tiếp theo để khắc phục sự khác nhau giữa hai ước lượng đói nghèo. Một cách tiếp cận là nên chọn những huyện có hai ước lượng khác nhau rõ ràng (ví dụ như hai trường hợp ở trên) và thu thập các số liệu thứ cấp và sơ cấp để xác định ước lượng nào là chính xác với thực tế nhất.

Hình 23. So sánh tỷ lệ đói nghèo (P0) từ MOLISA và từ phương pháp ước lượng diện tích nhỏ



4. Các yếu tố địa lý ảnh hưởng đến đói nghèo

Các bản đồ đói nghèo ở trên cho thấy sự biến đổi đáng kể về mặt địa lý giữa các tỉnh, huyện và xã ở Việt Nam. Tỷ lệ đói nghèo đặc biệt cao ở các vùng miền núi giáp biên giới với Trung Quốc và Lào, tỷ lệ đói nghèo thấp nhất ở các trung tâm đô thị lớn, ở khu vực đồng bằng châu thổ sông Hồng và sông Cửu Long. Phần này sử dụng ước lượng đói nghèo cấp huyện từ Phần 3 để kiểm tra mức độ ảnh hưởng của các biến địa lý tới tỷ lệ đói nghèo ở một huyện.

4.1. Các yếu tố địa lý

Bảng 9 là danh sách các biến địa lý có thể giúp các nhà nghiên cứu giải thích sự tác động của yếu tố không gian tới tỷ lệ đói nghèo ở Việt Nam. Các biến này được chia làm 2 nhóm. Biến ngoại sinh là những biến không bị ảnh hưởng bởi mức độ hoạt động kinh tế hoặc đói nghèo. Ví dụ, các biến khí hậu nông nghiệp như lượng mưa hay địa hình có thể ảnh hưởng đến đói nghèo, song chúng không thể bị ảnh hưởng bởi đói nghèo. Trái lại, biến nội sinh có thể vừa ảnh hưởng tới đói nghèo vừa bị ảnh hưởng bởi nó (ít nhất là trong thời gian dài). Ví dụ, các vùng có tỷ lệ đói nghèo thấp có thể thu hút người nhập cư làm dân số trong vùng tăng lên. Tương tự như vậy, đầu tư cho chợ và cơ sở hạ tầng giao thông cũng bị ảnh hưởng một phần nào bởi mức độ hoạt động kinh tế, do đó tỷ lệ đói nghèo thấp có thể ảnh hưởng tới mật độ chợ và đường xá trong dài hạn.

Bảng 9. Các yếu tố khí hậu nông nghiệp và kinh tế xã hội có thể ảnh hưởng tới tỷ lệ đói nghèo

Tên biến	Mối quan hệ kỳ vọng với đói nghèo
Biến ngoại sinh	
Độ cao	Độ cao lớn → tỷ lệ đói nghèo cao
Độ dốc/Độ gồ ghề	Độ dốc lớn → tỷ lệ đói nghèo cao
Loại đất	Đất cát và cằn → tỷ lệ đói nghèo cao
Đất mặt	Không rõ
Số giờ nắng	Ít được chiếu sáng → tỷ lệ đói nghèo cao
Lượng mưa hàng năm	Lượng mưa ít → tỷ lệ đói nghèo cao
Khoảng cách tới thị xã và thành phố	Khoảng cách xa → tỷ lệ đói nghèo cao
Biến nội sinh	
Dân số	Không rõ
Số lượng và mật độ chợ	Mật độ chợ thấp ↔ tỷ lệ đói nghèo cao
Chiều dài và mật độ đường xá	Mật độ đường xá thấp ↔ tỷ lệ đói nghèo cao
Chiều dài và mật độ sông	Mật độ sông thấp ↔ tỷ lệ đói nghèo cao
Thời gian ra tới thị xã và thành phố	Thời gian dài ↔ tỷ lệ đói nghèo cao

Cột bên phải của Bảng 9 cho thấy mối liên hệ dự đoán giữa mỗi biến và sự đói nghèo. Mũi tên hai chiều biểu thị các trường hợp mà trong đó đói nghèo và biến có tác động lẫn nhau.

Để thực hiện một phép phân tích hồi quy, các yếu tố khí hậu nông nghiệp trong Bảng 9 phải được coi như một biến cụ thể. Bảng 10 là danh sách các biến khí hậu nông nghiệp đã được kiểm tra trong phép phân tích này.

Bảng 10. Các biến được sử dụng để phân tích các yếu tố địa lý

Tên biến	Mô tả
e_251_500m	Tỉ lệ tỉnh có độ cao từ 251 đến 500 m
e_501_1000m	Tỉ lệ tỉnh có độ cao từ 501 đến 1000 m
e_1001_1500m	Tỉ lệ tỉnh có độ cao từ 1001 đến 1500 m
e_over_1500m	Tỉ lệ tỉnh có độ cao trên 1500 m
lc_arable	Đất mặt: Tỉ lệ đất có thể canh tác
lc_barerocky	Đất mặt: Tỉ lệ đất cằn và đất đá
lc_natfor	Đất mặt: Tỉ lệ rừng quốc gia
lc_plantfor	Đất mặt: Tỉ lệ rừng trồng
pctslope2	Độ dốc: Tỉ lệ đất có độ dốc từ 4 - 8%
pctslope3	Độ dốc: Tỉ lệ đất có độ dốc từ 8 - 15%
pctslope4	Độ dốc: Tỉ lệ đất có độ dốc từ 15 - 30%
pctslope5	Độ dốc: Tỉ lệ đất có độ dốc trên 30%
s_alluv	Tỉ lệ tỉnh có đất phù sa
s_alluvglay	Tỉ lệ tỉnh có đất phù sa băng
s_alluvacid	Tỉ lệ tỉnh có đất phù sa axit
s_acidsulf	Tỉ lệ tỉnh có đất muối
s_avgsalt	Tỉ lệ tỉnh có đất mặn
s_alluvrust	Percent of district area with Alluvial Rust
s_redbrown	Percent of district area with Red Brown Sols
s_sandy	Percent of district area with Sandy Soils
s_othfluv	Percent of district area with Fluvial Soils
s_othacris	Tỉ lệ tỉnh có đất Acrisol Soils
s_other	Tỉ lệ tỉnh có các loại đất khác
s_rocky	Tỉ lệ tỉnh có đất đá
s_water	Tỉ lệ tỉnh có sông, hồ và ao
sun_annual	Thời gian được chiếu sáng hàng năm (ngày)
prec_annual	Lượng mưa hàng năm (mm)
d_dist_10k	Khoảng cách từ thị xã với hơn 10.000 người
d_dist_50k	Khoảng cách từ thị xã với hơn 50.000 người
d_dist_100k	Khoảng cách từ thị xã với hơn 100.000 người
d_dist_250k	Khoảng cách từ thị xã với hơn 250.000 người
d_dist_1m	Khoảng cách từ thị xã với hơn 1 triệu người

Bằng phép kiểm định Chow, chúng ta thấy rằng các hệ số giải thích mô hình đói nghèo cho nông thôn và thành thị khác nhau đáng kể. Điều này cho thấy việc tách biệt các mô hình nông thôn và thành thị sẽ thích hợp hơn. Chúng ta sẽ giả thiết loại đất, đất mặt, lượng mưa và số ngày nắng quan trọng hơn ở các vùng nông thôn, nơi có tới 2/3 số hộ gia đình coi nông nghiệp là hoạt động kinh tế chính. Trái lại, các chỉ số của hoạt động phi nông nghiệp như là số chợ trên mỗi tỉnh và khoảng cách tới các thành phố lớn sẽ trở nên quan trọng hơn ở thành thị.

4.2. Các vấn đề trong ước lượng

Tỉ lệ đói nghèo ở các huyện lân cận có nhiều khả năng giống nhau, do đó điều quan trọng là phải chú ý tới cấu trúc của sự phụ thuộc về mặt không gian trong số liệu đã phân tích trong phần 2.6. Nếu không chú ý tới cấu trúc này, những ước lượng về sự ảnh hưởng của các biến địa lý có thể không thống nhất và bị chệch, nhất là khi dùng phương pháp bình phương nhỏ nhất thông thường (ordinary least squares) để ước lượng. Những tài liệu về kinh tế lượng không gian phân biệt 2 loại phụ thuộc về không gian như sau:

- i. Phụ thuộc sai số về không gian (spatial error dependence), trong đó các biến giải thích không được nghiên cứu có mối liên hệ về không gian. Ví dụ, do chính sách và ngân sách của tỉnh, chất lượng y tế địa phương ở các huyện trong một tỉnh giống nhau, song lại giữa các tỉnh lại khác nhau. Khi có phụ thuộc sai số về không gian, hệ số hồi quy của phương pháp bình phương nhỏ nhất thông thường (ordinary least squares) sẽ bị chệch (biased) song cũng không hiệu quả (sai số chuẩn sẽ lớn hơn so với khi sử dụng tất cả thông tin).
- ii. Phụ thuộc trễ về không gian (spatial lag dependence), trong đó biến phụ thuộc trong một vùng bị ảnh hưởng trực tiếp bởi các biến phụ thuộc ở các vùng lân cận. Ví dụ, tỉ lệ đói nghèo ở một huyện bị ảnh hưởng trực tiếp bởi đói nghèo ở các huyện lân cận. Khi có phụ thuộc trễ về không gian (spatial lag dependence), hệ số hồi quy của phương pháp bình phương nhỏ nhất bình thường (ordinary least squares) sẽ bị chệch và không thống nhất.

Bất cứ khi nào có sự phụ thuộc sai số về không gian (spatial error dependence) hoặc sự phụ thuộc trễ về không gian (spatial lag dependence), cần phải áp dụng các kiểu mô hình hồi quy đặc biệt. Trong trường hợp có sự phụ thuộc sai số về không gian (spatial error dependence), mô hình sai số không gian là chính xác, và trong trường hợp có sự phụ thuộc trễ về không gian (spatial lag dependence), mô hình trễ về không gian sẽ được sử dụng. Các bước ước lượng của chúng tôi như sau. Đầu tiên, chúng tôi ước lượng một mô hình bình phương nhỏ nhất bình thường (ordinary least squares) (OLS) bao gồm tất cả các biến ngoại sinh. Thứ hai, chúng tôi kiểm tra hai kiểu phụ thuộc không gian. Thứ ba, chúng tôi sẽ sử dụng mô hình sai số về không gian (spatial error) hoặc là trễ về không gian (spatial lag) để ước lượng lại mô hình. Có hai phiên bản cho mô hình đói nghèo cho nông thôn và thành thị : một kiểu chọn lọc

hơn chỉ bao gồm các biến ngoại sinh, một kiểu toàn diện hơn gồm một vài biến có thể là nội sinh, ít nhất trong dài hạn.

4.3. Mô hình tổng thể phân tích đói nghèo nông thôn

Bảng 11 cho biết các kiểm định sự phụ thuộc về không gian khi một mô hình bình phương nhỏ nhất bình thường (ordinary least squares) được ước lượng với tỉ lệ đói nghèo nông thôn đến cấp huyện như biến phụ thuộc và các biến ngoại sinh đã phân tích ở trên. Quyền số khoảng cách nghịch đảo (inverse distance weights) được sử dụng để biểu diễn kiểm tra này. Cả hệ số I của Moran và hệ số Lagrange dùng để kiểm tra kết quả đều phủ nhận giả thuyết H_0 về không có sự phụ thuộc về không gian nào. Hệ số Lagrange trong mô hình càng lớn cho biết có sự phụ thuộc về không gian. Vì thế, chúng tôi tiến hành ước lượng mô hình sai số không gian để phân tích các yếu tố ảnh hưởng đến đói nghèo ở nông thôn.

Bảng 11. Kiểm tra sự phụ thuộc không gian trong mô hình đói nghèo cho nông thôn

Kiểm tra	Kết quả	df	giá trị p
Mô hình sai số về không gian:			
Hệ số I của Moran	25,459	1	0,000
Hệ số Robust Lagrange	168,329	1	0,000
Mô hình trễ về không gian:			
Hệ số Robust Lagrange	6,166	1	0,013

Mô hình đầy đủ đói nghèo cho nông thôn (inclusive model of rural poverty)

Bảng 12 là kết quả của phép hồi quy tỉ lệ đói nghèo nông thôn cấp huyện với tập hợp đầy đủ các biến ngoại sinh sử dụng mô hình sai số về không gian. Mô hình này giải thích được 4/5 sự biến động về tỉ lệ đói nghèo nông thôn, song chỉ có 8 trong số 32 hệ số là có ý nghĩa thống kê với độ tin cậy 5%. Điều ngạc nhiên là không có biến về độ cao nào có ý nghĩa thống kê, song điều này hoàn toàn có thể xảy ra bởi độ dốc và loại đất cũng có mặt trong mô hình này. Nói cách khác, độ cao không ảnh hưởng trực tiếp tới đói nghèo ở nông thôn, ngoại trừ khi chúng được kết hợp với đất đai cằn cỗi và độ dốc lớn.

Trong số các biến đất mặt, chỉ có đất cằn và đất đá có ý nghĩa thống kê. Hệ số dương có nghĩa là tỉ lệ đói nghèo cao hơn ở các vùng có tỉ lệ đất cằn và đất đá cao.

Bảng 12. Mô hình tổng thể các yếu tố địa lý ảnh hưởng tới đói nghèo ở nông thôn

Mô hình sai số về không gian	Number of obs = 569					
	Variance ratio = 0.679					
	Squared corr. = 0.736					
Log likelihood = 695.58012	Sigma = 0.07					
R-squared = 0.8009						
	Robust					
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
p0rural						
e_251_500m	.0005366	.0004063	1.32	0.187	-.0002598	.001333
e_501_1000m	.0004463	.0003416	1.31	0.191	-.0002233	.001116
e_1001_1500m	-.0002176	.0008015	-0.27	0.786	-.0017885	.0013533
e_over_1500m	.0012043	.0010897	1.11	0.269	-.0009313	.00334
lc_arable	.0003025	.0001767	1.71	0.087	-.0000438	.0006488
lc_barerocky	.0016445	.0005611	2.93	0.003	.0005447	.0027443
lc_natfor	-.0000316	.0002933	-0.11	0.914	-.0006064	.0005431
lc_plantfor	-.0017524	.0010513	-1.67	0.096	-.0038128	.0003081
pctslope2	.0039081	.0008831	4.43	0.000	.0021772	.0056389
pctslope3	.0010364	.0012501	0.83	0.407	-.0014138	.0034866
pctslope4	.0039068	.0011927	3.28	0.001	.0015691	.0062445
pctslope5	.0034696	.001971	1.76	0.078	-.0003935	.0073327
s_alluv	-.000042	.0002639	-0.16	0.873	-.0005592	.0004752
s_alluvglay	.0003011	.0003052	0.99	0.324	-.0002971	.0008993
s_alluvacid	.0000304	.000292	0.10	0.917	-.000542	.0006027
s_acidsulf	.000612	.0003086	1.98	0.047	7.12e-06	.001217
s_avgsalt	.0005966	.0002946	2.03	0.043	.0000192	.0011741
s_alluvrust	-.0001836	.0003344	-0.55	0.583	-.0008391	.0004719
s_redbrown	-.0007071	.0003803	-1.86	0.063	-.0014525	.0000383
s_sandy	.0012598	.0004398	2.86	0.004	.0003979	.0021218
s_othfluv	.0006494	.0003864	1.68	0.093	-.0001079	.0014068
s_othacris	.0009033	.000417	2.17	0.030	.000086	.0017205
s_other	-.0000894	.0003413	-0.26	0.793	-.0007585	.0005796
s_rocky	-.0009221	.0006002	-1.54	0.124	-.0020984	.0002543
s_water	.0004779	.0009002	0.53	0.596	-.0012865	.0022423
sun_annual	-.000037	.0000544	-0.68	0.497	-.0001436	.0000697
prec_annual	-8.81e-06	.0000312	-0.28	0.778	-.00007	.0000524
d_dist_10k	2.50e-06	3.45e-07	7.26	0.000	1.83e-06	3.18e-06
d_dist_50k	5.95e-07	3.31e-07	1.80	0.072	-5.35e-08	1.24e-06
d_dist_100k	2.53e-07	2.89e-07	0.88	0.380	-3.12e-07	8.19e-07
d_dist_250k	8.20e-08	2.99e-07	0.27	0.784	-5.04e-07	6.68e-07
d_dist_1m	-2.30e-08	3.03e-07	-0.08	0.939	-6.16e-07	5.70e-07
_cons	.3901486	.1326071	2.94	0.003	.1302435	.6500538
lambda	.9015347	.0421281	21.40	0.000	.8189651	.9841043
Wald test of lambda=0: chi2(1) = 457.953 (0.000)						
Kiểm định nhân tử Lagrange của lambda=0: chi2(1) = 378.471 (0.000)						
khoảng chấp nhận cho lambda: -1.900 < lambda < 1.000						

Hai biến độ dốc có ý nghĩa thống kê là tỉ lệ đất có độ dốc từ 4-8% và tỉ lệ đất có độ dốc từ 15-30%. Do bỏ qua biến đất bằng phẳng (không dốc), các kết quả này cho thấy các huyện có diện tích đất dốc lớn sẽ có tỉ lệ đói nghèo cao hơn các tỉnh có diện tích đất bằng phẳng. Điều này không có gì ngạc nhiên bởi địa hình đất dốc sẽ bất lợi cho trồng trọt và tưới tiêu.

Các hệ số của ba loại đất cũng có ý nghĩa thống kê với độ tin cậy 5% (do đất sắt xám là loại đất phổ biến nhất, chúng tôi đã sử dụng như loại biến khuyết). Ví dụ, các huyện có tỉ lệ đất mặn là đất nhiễm muối axit lớn có xu hướng nghèo hơn các huyện có tỉ lệ đất loại này thấp hơn. Các huyện có diện tích đất mặn lớn cũng thường nghèo hơn các huyện có ít diện tích đất mặn và cũng tương tự như vậy với những huyện có diện tích đất cát lớn. Đất nhiễm muối axit là vấn đề đặc biệt nghiêm trọng ở phần lớn các tỉnh Đồng bằng sông Cửu Long. Muối ảnh hưởng tới diện tích đất nông nghiệp được tưới gần các vùng duyên hải không có đủ cơ sở hạ tầng hoặc nước sạch để tránh bị nhiễm mặn bởi nước biển. So với các vùng khác, đất cát lại là trở ngại lớn nhất ở vùng Duyên hải Nam Trung bộ.

Điều ngạc nhiên là biến đại diện lượng mưa trung bình hàng năm và số ngày nắng không có ý nghĩa thống kê đối với tỉ lệ đói nghèo ở nông thôn. Điều này có thể hiểu được bởi ở những nơi đất được tưới, lượng mưa thường ít hơn. Có lẽ ở những nơi được tưới tự nhiên, những vấn đề về đất dốc và đất cằn sẽ có ảnh hưởng lớn hơn so với sự thay đổi về lượng mưa và ánh nắng mặt trời.

Trong số các biến đại diện cho khoảng cách tới trung tâm đô thị lớn, khoảng cách tới một thành phố ít nhất 10.000 người là một hệ số dương và có ý nghĩa thống kê. Điều này có nghĩa là, nếu các biến khác không đổi, một vùng cách xa thị xã khoảng 10.000 người có xu hướng nghèo hơn các vùng gần thị xã hơn. Khoảng cách tới một thị xã là 50.000 người không có ý nghĩa thống kê với độ tin cậy 5%, song lại có ý nghĩa với độ tin cậy 10%. Điều đáng ngạc nhiên là khoảng cách tới các thành phố lớn lại không có ý nghĩa thống kê.

Mô hình chọn lọc đói nghèo ở nông thôn (selective model of rural poverty)

Bây giờ, chúng tôi sẽ chuyển sang nghiên cứu mô hình chọn lọc đói nghèo cho nông thôn thông qua việc kết hợp trực giác kinh tế và hồi quy bậc thang (stepwise backward regression) để loại bỏ hoặc kết hợp các biến từ mô hình đầy đủ. Do lượng mưa và số ngày nắng là những chỉ số ước lượng không có ý nghĩa đối với đói nghèo ở nông thôn trong mô hình sai số về không gian (spatial error model) nên chúng bị loại ra. Sau đó, chúng tôi đã tiến hành các kiểm tra F-test và thấy hiệu quả khi kết hợp tất cả bốn biến đất phù sa với các loại đất trong nhóm loại bỏ (omitted category) đất sắt xám (loại đất chiếm khoảng 37% diện tích đất ở Việt Nam). Bên cạnh đó, do Việt Nam có rất ít diện tích đất có độ dốc trên 15% hoặc độ cao trên 500 m so với mặt nước biển, chúng tôi đã kết hợp các nhóm độ dốc và độ cao với các biến gần nhất. Cuối cùng, chúng tôi tập trung nghiên cứu khoảng cách tới thị xã với số dân trên 10.000 và 100.000 người và các thành phố trên 1 triệu người nhằm giảm bớt khả năng đa cộng tuyến (multicollinearity) với các biến khoảng cách khác.

Kết quả của mô hình phân tích hồi quy này được mô tả trong Bảng 13 với 12 biến giải thích. Mô hình này vẫn có khả năng giải thích 74% sự biến động về tỉ lệ đói nghèo, với tỉ lệ đói nghèo tăng ở những vùng đất dốc hoặc đất có độ axit cao, đất muối hoặc các loại đất chua khác. Đất mặt cằn cỗi và đất đá cũng sẽ làm cho tỉ lệ đói nghèo ở nông thôn tăng lên, mặc dù ảnh hưởng của độ cao (yếu tố có liên quan với độ dốc, loại đất và đất mặt) không có ý nghĩa. Cuối cùng, khoảng cách tới các thị xã nhỏ và trung bình (trên 100.000 người) có tương quan cùng chiều với tỷ đói nghèo ở nông thôn, song khoảng cách tới các thành phố trên 1 triệu dân thì không có ảnh hưởng. Kết quả này có thể giả thích khi có dẫn chứng về tầm quan trọng của việc gần các thị xã nhỏ hơn là các thành phố lớn trong việc giảm bớt tỉ lệ đói nghèo.

Bảng 13. Mô hình chọn lọc các yếu tố địa lý ảnh hưởng đến đói nghèo

Mô hình sai số không gian			Number of obs	=	569	
			Variance ratio	=	0.639	
			Squared corr.	=	0.723	
Log likelihood = 676.05323			Sigma	=	0.07	
R-squared = 0.7379						
	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
p0rural						
pctslope2	.0023138	.000644	3.59	0.000	.0010515	.0035761
pctslope34	.0025204	.0004661	5.41	0.000	.0016069	.0034339
pctslope5	.0047158	.0014319	3.29	0.001	.0019093	.0075223
s_acidsulf	.0006972	.0002769	2.52	0.012	.0001546	.0012398
s_avgsalt	.0006522	.0001998	3.26	0.001	.0002606	.0010438
s_othacris	.0007839	.0003169	2.47	0.013	.0001628	.0014051
lc_barerocky	.001079	.000394	2.74	0.006	.0003068	.0018513
d_dist_10k	2.75e-06	3.26e-07	8.45	0.000	2.11e-06	3.39e-06
d_dist_100k	6.09e-07	2.65e-07	2.29	0.022	8.86e-08	1.13e-06
d_dist_1m	4.82e-08	1.96e-07	0.25	0.806	-3.37e-07	4.33e-07
e_25l_500m	.0003639	.000395	0.92	0.357	-.0004103	.0011382
e_over_500m	.000278	.0003297	0.84	0.399	-.0003682	.0009243
_cons	.3247383	.0408911	7.94	0.000	.2445932	.4048834
lambda	.8972043	.0345779	25.95	0.000	.8294329	.9649758
Wald test of lambda=0:			chi2(1) = 673.264 (0.000)			
Lagrange multiplier test of lambda=0:			chi2(1) = 961.892 (0.000)			
Acceptable range for lambda: -1.900 < lambda < 1.000						

Với mô hình chọn lọc này, chúng tôi đã tiến hành kiểm tra xem việc kết hợp các biến nội sinh có thể làm tăng khả năng giải thích của mô hình hay không? Về lý thuyết, việc thêm vào các biến nội sinh có thể gây ảnh hưởng về tính đồng nhất (simultaneity¹⁷). Tuy nhiên, trên thực

¹⁷ Nếu các biến giải thích trong phương trình hồi quy bị ảnh hưởng bởi biến độc lập thì hệ số ước lượng bởi phương bình phương nhỏ nhất bình thường có thể bị chệch và không phản ánh được sự thay đổi của biến giải thích tới biến độc lập

tế, việc thêm vào các biến nội sinh như dân số và mật độ đường xá trong mô hình đói nghèo cho nông thôn chỉ làm tăng 3% khả năng giải thích (R^2). Bằng cách phân loại mô hình khác khi thay thế khoảng cách tới thị xã và thành phố bằng thời gian di chuyển, và mật độ dân trong vùng bằng mật độ chợ, cũng cho kết quả tương tự.

Cuối cùng, khi dùng các biến giả cho 7 vùng ở Việt Nam trong mô hình đói nghèo cho nông thôn, chúng có ý nghĩa như nhau và làm tăng khả năng giải thích của mô hình 9%. Tuy nhiên, trong nhiều trường hợp, các biến giả lại làm tăng ảnh hưởng của các biến địa lý bị loại bỏ và được giải thích tốt nhất như là những tác động cố định cho thấy sự mất khả năng trong việc giải thích từ 75 – 80% sự biến động về tỉ lệ đói nghèo ở nông thôn bằng các biến địa lý.

Chúng tôi không hề ngạc nhiên khi thấy rằng các biến khí hậu nông nghiệp có ảnh hưởng tới đói nghèo ở nông thôn. Tỉ lệ đói nghèo ở các vùng nông thôn có quan hệ mật thiết với năng suất nông nghiệp và khả năng tiếp cận thị trường. Đất mặt, độ dốc và loại đất có ảnh hưởng trực tiếp với năng suất nông nghiệp. Tương tự như vậy, khoảng cách tới thị xã và thành phố là một chỉ số đánh giá khả năng tiếp cận thị trường, chỉ số này có ảnh hưởng tới giá cả hàng hoá nông dân nhận được cũng như giá họ phải trả cho vật tư đầu vào. Tuy nhiên, điều ngạc nhiên là các biến này giải thích được nhiều sự biến động về tỉ lệ đói nghèo giữa các vùng.

4.4. Mô hình tổng thể đói nghèo thành thị (global model of urban poverty)

Giống như mô hình đói nghèo cho nông thôn, đầu tiên chúng tôi kiểm tra kiểu phụ thuộc về không gian khi một mô hình bình phương nhỏ nhất bình thường không giới hạn (unrestricted ordinary least squares) được ước lượng với tỉ lệ đói nghèo ở thành thị đến cấp huyện như một biến độc lập và các biến ngoại sinh ở trên. Bảng 14 là kiểm định sự phụ thuộc về không gian (quyền số khoảng cách nghịch đảo được sử dụng để thực hiện việc kiểm tra này). Các kết quả hiện nay ít chính xác hơn so với trước đây. Đối với mô hình tổng thể cho đói nghèo thành thị gồm tất cả các biến ngoại sinh, có rất ít bằng chứng cho thấy ưu thế của sự phụ thuộc về không gian hay là sự phụ thuộc sai số về không gian. Điều này có thể được giải thích khi đưa ra một số dẫn chứng về tỉ lệ đói nghèo ở một huyện thành thị bị ảnh hưởng trực tiếp bởi tỉ lệ đói nghèo ở các huyện lân cận. Tuy nhiên, việc kiểm định này cũng cho thấy mô hình sai số về không gian (spatial error model) có thể thích hợp và hiệu quả của mô hình này được khẳng định một cách độc lập so với mô hình chọn lọc (có hoặc không có biến nội sinh) như trình bày dưới đây.

Bảng 14: Kiểm định sự phụ thuộc về không gian đối với mô hình đói nghèo thành thị

Test	Statistic	df	p-value
Spatial error:			
Moran's I	18.674	1	0.000
Lagrange multiplier	192.773	1	0.000
Robust Lagrange multiplier	11.678	1	0.001
Spatial lag:			
Lagrange multiplier	195.959	1	0.000
Robust Lagrange multiplier	14.864	1	0.000

Mô hình đầy đủ cho đói nghèo thành thị (inclusive model)

Giống như ở trên, chúng tôi bắt đầu với một mô hình đủ và sau đó xác định một mô hình cụ thể hơn. Mô hình đầy đủ cho đói nghèo thành thị đến cấp huyện có thể giải thích 38% sự biến động về tỉ lệ đói nghèo ở thành thị (xem Bảng 15). Điều này chứng tỏ rằng đói nghèo ở thành thị khó giải thích với các biến địa lý hơn so với đói nghèo ở nông thôn.

Chỉ có 4 trong số 32 biến giải thích có hệ số có ý nghĩa thống kê với độ tin cậy 5%. Tỉ lệ đất có thể canh tác có liên quan chặt chẽ đến đói nghèo ở thành thị, song với tương quan cùng chiều, cho thấy diện tích đất canh tác càng lớn sẽ cho tỉ lệ đói nghèo càng cao. Có thể khu vực thành thị ở những huyện có nhiều đất nông nghiệp là các thị xã nhỏ, do đó biến này có thể làm tăng ảnh hưởng của quy mô thị xã tới đói nghèo.

Bên cạnh đó, đói nghèo thành thị có liên quan tới tỉ lệ vùng có đất đỏ bazan (red-brown soil). Rất khó có thể giải thích kết quả này bởi đơn giản nó có thể phản ánh một sự tương quan giả. Đói nghèo thành thị cũng có liên quan tới khoảng cách tới vùng gần nhất với hơn 10.000 người và khoảng cách tới thị xã gần nhất với hơn 100.000 người.

Bảng 15. Mô hình đầy đủ các yếu tố địa lý ảnh hưởng đến đói nghèo thành thị

Spatial error model			Number of obs	=	574	
			Variance ratio	=	0.358	
			Squared corr.	=	0.289	
Log likelihood = 865.43392			Sigma	=	0.05	
R-squared	=	0.3783				
	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
p0urban						
e_251_500m	.0001082	.0002731	0.40	0.692	-.0004271	.0006435
e_501_1000m	-.0000278	.0002845	-0.10	0.922	-.0005853	.0005298
e_1001_1500m	-.000056	.0005433	-0.10	0.918	-.0011209	.0010088
e_over_1500m	-.0000753	.00111	-0.07	0.946	-.0022508	.0021002
lc_arable	.0003578	.0000948	3.77	0.000	.000172	.0005436
lc_barerocky	.0003138	.0005151	0.61	0.542	-.0006959	.0013234
lc_natfor	.0001816	.0001717	1.06	0.290	-.000155	.0005181
lc_plantfor	-.0010288	.0006689	-1.54	0.124	-.0023397	.0002822
pctslope2	.0005427	.0006581	0.82	0.410	-.0007471	.0018325
pctslope3	.0007459	.0009565	0.78	0.436	-.0011289	.0026206
pctslope4	.0009161	.0008757	1.05	0.295	-.0008002	.0026324
pctslope5	.0007792	.0015636	0.50	0.618	-.0022853	.0038437
s_alluv	-.0001124	.0001845	-0.61	0.542	-.000474	.0002492
s_alluvglay	.0001253	.0002099	0.60	0.550	-.000286	.0005367
s_alluvacid	.0000288	.0002096	0.14	0.891	-.000382	.0004396
s_acidsulf	.0001295	.000201	0.64	0.519	-.0002644	.0005235
s_avgsalt	.0001363	.0001762	0.77	0.439	-.0002091	.0004817
s_alluvrust	-.0003464	.0002077	-1.67	0.095	-.0007534	.0000606
s_redbrown	-.0007689	.0002382	-3.23	0.001	-.0012357	-.0003021

s_sandy	.0004235	.0003504	1.21	0.227	-.0002633	.0011103
s_othfluv	.0001111	.0002161	0.51	0.607	-.0003125	.0005347
s_othacris	.0003841	.0002917	1.32	0.188	-.0001876	.0009559
s_other	.0001843	.0002943	0.63	0.531	-.0003926	.0007612
s_rocky	-.0000639	.0005357	-0.12	0.905	-.0011138	.000986
s_water	.0001654	.000528	0.31	0.754	-.0008694	.0012003
sun_annual	.0000102	.0000239	0.43	0.669	-.0000366	.000057
prec_annual	-6.14e-06	.0000206	-0.30	0.766	-.0000466	.0000343
d_dist_10k	5.74e-07	2.42e-07	2.37	0.018	9.94e-08	1.05e-06
d_dist_50k	5.11e-07	2.12e-07	2.41	0.016	9.60e-08	9.27e-07
d_dist_100k	-1.40e-07	1.79e-07	-0.78	0.434	-4.90e-07	2.10e-07
d_dist_250k	-1.12e-10	1.42e-07	-0.00	0.999	-2.79e-07	2.79e-07
d_dist_1m	1.17e-07	1.02e-07	1.14	0.252	-8.36e-08	3.18e-07
_cons	.0992521	.0687563	1.44	0.149	-.0355078	.2340119
lambda	.8127389	.0520617	15.61	0.000	.7106999	.914778
Wald test of lambda=0:			chi2(1) = 243.706 (0.000)			
Lagrange multiplier test of lambda=0:			chi2(1) = 192.773 (0.000)			
Khoảng chấp nhận của lambda: -1.900 < lambda < 1.000						

Mô hình chọn lọc đánh giá đói nghèo thành thị

Một mô hình chọn lọc phân tích các yếu tố ảnh hưởng đến đói nghèo cho thành thị được phát triển bằng cách sử dụng các bước tương tự như đối với mô hình chọn lọc cho đói nghèo cho nông thôn. Tất cả các loại đất trừ đất đỏ bazan (red-brown) được loại bỏ hoàn toàn khi sử dụng phương pháp hồi quy bậc thang (stepwise regression). Độ cao tiếp tục được coi là một biến có ý nghĩa thống kê ảnh hưởng đến đói nghèo. Tuy nhiên, ở các khu vực thành thị đông dân, đây không phải là tỉ lệ đất trong các lớp độ cao khác nhau. Do đó, các trung tâm đô thị lớn ở vùng đồng bằng châu thổ và ven biển có độ biến động về độ cao rất nhỏ, biến này cho biết các trung tâm đô thị ở vùng trung du và cao nguyên thường nghèo hơn ở các vùng duyên hải và đồng bằng châu thổ. Cuối cùng, sau một vài phép thử, mô hình đói nghèo cho thành thị vẫn giữ lại ba biến đại diện khoảng cách tới thị xã và thành phố. Tuy nhiên, khác với mô hình cho nông thôn, khoảng cách tới các thành phố lớn và trung bình quan trọng hơn. Điều này cho thấy tầm quan trọng của việc sống gần các trung tâm đô thị lớn, đặc biệt ở Hà Nội và thành phố Hồ Chí Minh, hai thành phố duy nhất có dân số hơn 1 triệu người, nhất là đối với những người sản xuất hàng hoá phi nông nghiệp và dịch vụ.

Cần lưu ý rằng phiên bản cuối cùng của mô hình chọn lọc phân tích đói nghèo ở thành thị được ước lượng bằng mô hình sai số không gian, khi chạy lại các kiểm định (diagnostic tests) sự phụ thuộc về không gian (spatial dependence) cho thấy mô hình sai số không gian hiện được sử dụng nhiều hơn (hệ số Lagrange bằng 371.710 đối với mô hình sai số không gian so với 293,1 đối với mô hình phụ thuộc trễ về không gian).

Bảng 16 . Mô hình lựa chọn phân tích các yếu tố địa lý ảnh hưởng tới đói nghèo thành thị

Mô hình sai số không gian	số quan sát	=	573
	Phương sai	=	0.339
	Squared corr.	=	0.254
Log likelihood = 841.82739	Sigma	=	0.05
R-squared	=	0.2917	

		Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
p0urban							
elev_town		.0000132	.0000229	0.58	0.564	-.0000317	.0000581
elev_stdv		.0001964	.0000361	5.45	0.000	.0001257	.000267
s_redbrown		-.0009041	.0002049	-4.41	0.000	-.0013057	-.0005025
d_dist_100k		2.39e-07	1.71e-07	1.40	0.161	-9.56e-08	5.74e-07
d_dist_250k		4.69e-08	1.41e-07	0.33	0.740	-2.30e-07	3.23e-07
d_dist_1m		7.52e-08	9.93e-08	0.76	0.449	-1.19e-07	2.70e-07
_cons		.1407774	.019244	7.32	0.000	.1030598	.178495
lambda		.8023299	.0504752	15.90	0.000	.7034003	.9012594

Wald test of lambda=0:	chi2(1) = 252.668 (0.000)
Lagrange multiplier test of lambda=0:	chi2(1) = 371.710 (0.000)

Acceptable range for lambda: -1.900 < lambda < 1.000

Tuy nhiên, cũng cần nhấn mạnh rằng khả năng giải thích của mô hình lựa chọn đói nghèo thành thị không lớn: 6 biến địa lý trong mô hình này chỉ giải thích ít hơn 30% sự biến động về tỉ lệ đói nghèo ở thành thị của các tỉnh. Thêm vào đó, khả năng giải thích của mô hình này cũng không được cải thiện nhiều lắm khi thêm vào các biến nội sinh. Thậm chí sau biến dân số, mật độ đường xá và mật độ chợ được thêm vào mô hình, khả năng giải thích của mô hình ở thành thị cũng chỉ tăng lên 41%. Các biến giả cho vùng cũng không thật sự quan trọng, chỉ làm tăng thêm 7% khả năng thích hợp (R^2) của mô hình chọn lọc.

Cũng không có gì đáng ngạc nhiên khi thấy điều kiện khí hậu nông nghiệp và khả năng tiếp cận thị trường không ảnh hưởng mấy tới các hoạt động dịch vụ và công nghiệp ở các thành thị. Ở đây, sự phát triển của các khu công nghiệp, cấu trúc việc làm và cơ sở hạ tầng sẽ là các yếu tố chủ yếu. Cần phải thêm vào một tập hợp biến khác, có thể là nội sinh, để giải thích các yếu tố địa lý có ảnh hưởng tới đói nghèo ở thành thị.

5. Biến động về không gian của các yếu tố ảnh hưởng tới đói nghèo

Phần 4 mô tả các kết quả phân tích của các mối liên hệ giữa sự phân bố về không gian của đói nghèo, các yếu tố sinh thái nông nghiệp và khả năng tiếp cận thị trường dựa trên một số mô hình hồi quy khác nhau. Các mô hình này có ý nghĩa "tổng thể" khi chúng mô tả mối liên hệ giữa đói nghèo và các yếu tố địa lý của Việt Nam.

Tuy nhiên, những mối liên hệ này có thể thay đổi giữa các vùng. Thật vậy, một đặc điểm của các số liệu về không gian là sự biến động trong không gian - một hiện tượng nhìn chung được xem là không dừng (non-stationarity) hay bị trôi đi (spatial drift). Vấn đề của non-stationarity có liên quan mật thiết với vấn đề đã mô tả ở trên đối với sự phụ thuộc về không gian : sai số của mô hình hồi quy tổng thể sẽ cho thấy sự tự tương quan không gian chúng ta sử dụng những số liệu có mối liên hệ về không gian, vì mô hình đói nghèo tổng thể chỉ có thể mô tả mối quan hệ tổng thể.

Phần này sẽ trình bày các kết quả của phân tích sự biến đổi về không gian những mối tương quan giữa tỉ lệ đói nghèo và một loạt các biến về sinh thái nông nghiệp. Chúng tôi sử dụng phép phân tích được gọi là phân tích hồi quy vùng có quyền số không gian (spatially-weighted local regression).

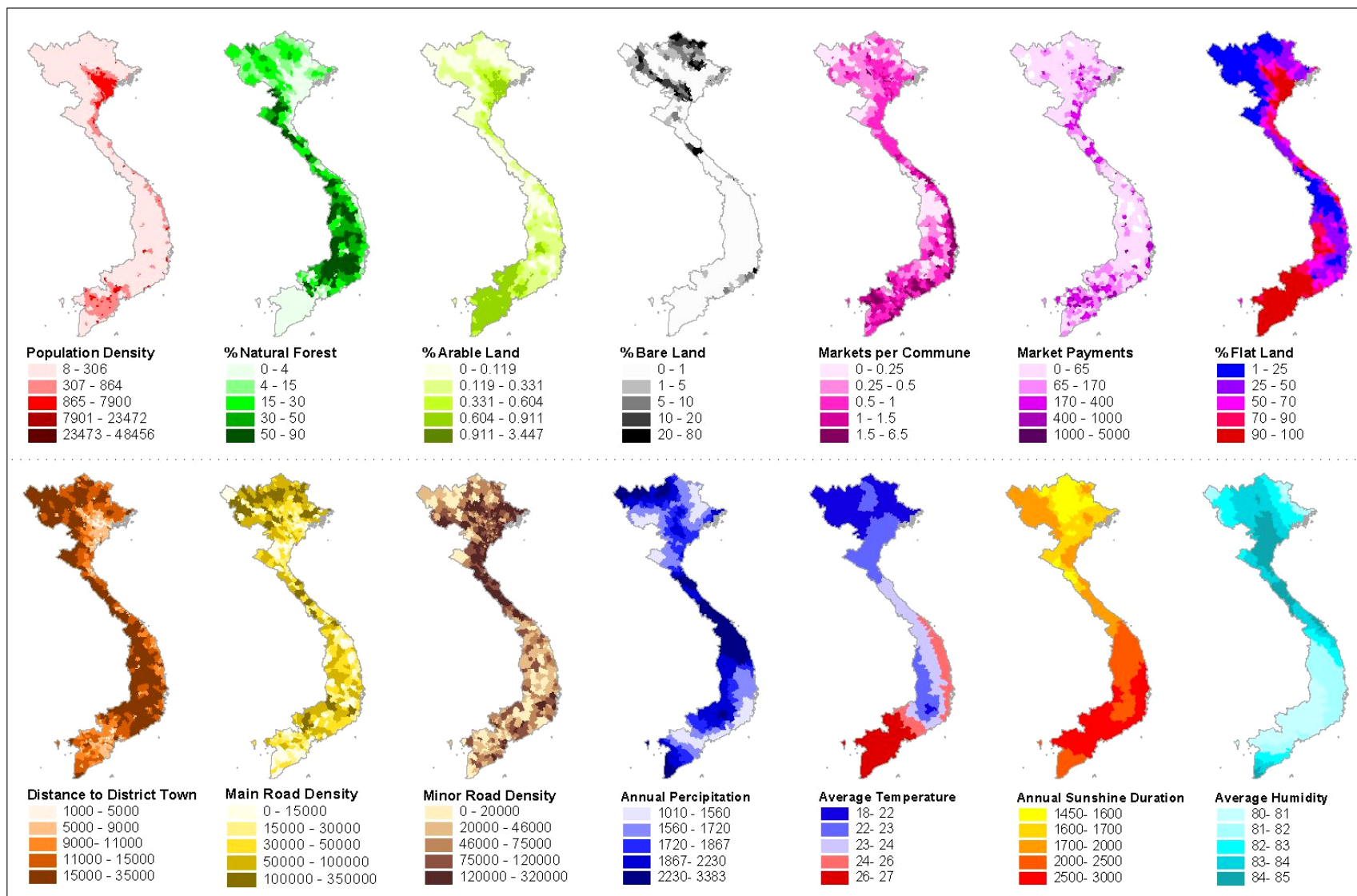
5.1. Mô tả mô hình

Tỷ lệ đói nghèo ước lượng cấp huyện P0 được xác định là một biến phụ thuộc, và 14 biến từ số liệu đã được mô tả trong phần 2.1 được lựa chọn làm các biến độc lập. 14 biến độc lập trong Bảng 17 và sự phân bố về mặt không gian của chúng được mô tả trong Hình 24.

Bảng 17. Các biến sử dụng trong phân tích hồi quy quyền số địa lý

Tên biến	Diễn giải
popdens	Mật độ dân số
lc natfo	% rừng tự nhiên
lc arabl	% diện tích có thể canh tác
lc barer	% đất cằn
mktperco	Số chợ trên một xã
mktpaym	thuế chợ nộp cho nhà nước
pctslope1	% đất bằng phẳng
d dist m	Khoảng cách trung bình ra thị xã
mainroad	Mật độ đường chính
minrroad	Mật độ đường nhỏ
prec ann	Lượng mưa hàng năm
temp avg	Nhiệt độ trung bình
sun annu	Thời gian được chiếu sáng trong năm
hum avg	Độ ẩm trung bình

Hình 24. Bản đồ phân bố không gian của các biến độc lập



Như đã mô tả trong phần 2.1, các biến được tính toán đến cấp huyện. Vị trí của các dữ liệu điểm mô tả các vị trí địa lý trong phân hồi vùng được xác định như trọng tâm của huyện.

Các điểm hồi quy được xác định là các trung tâm địa lý của mỗi huyện ở Việt Nam, phù hợp với các điểm số liệu đã định nghĩa ở trên. Các điểm hồi quy là các vị trí mà qua đó có thể ước lượng các tham số địa phương và từ đó các quan sát khác được xem xét với quyền số giảm dần.

5.2. Các kết quả

Mô hình hồi quy tổng thể được áp dụng đầu tiên cho các biến trước khi một mô hình hồi quy được "địa phương hoá-localized" bằng cách áp dụng phép hồi quy quyền số địa lý (geographically-weighted regression) . Bảng 18 biểu diễn kết quả của mô hình hồi quy tổng thể. Giá trị $R^2 = 0,75$ cho thấy sự phù hợp tương đối: 75% sự biến động tỉ lệ nghèo đến cấp huyện được giải thích. Gần như tất cả các biến đều có ý nghĩa (với độ tin cậy 1%), chỉ có duy nhất biến quy mô chợ (được tính toán bằng thuế chợ phải nộp cho nhà nước (payments of taxes from markets to the state) và tổng lượng mưa hàng năm dường như không liên quan tới biến phụ thuộc. Phần lớn các hệ số đều có chiều tương quan (dương hoặc âm). Ví dụ, mật độ đường xá sẽ có tác động ngược chiều với tỷ lệ đói nghèo, có nghĩa là tỉ lệ nghèo thấp khi mật độ đường xá cao. Điều duy nhất ngạc nhiên là mối quan hệ cùng chiều giữa tỉ lệ đất có thể canh tác và tỉ lệ đói nghèo, vì chúng ta hy vọng trong một nền kinh tế dựa phần lớn vào nông nghiệp, sự sẵn có của đất canh tác sẽ có ảnh hưởng tích cực tới phúc lợi mọi người.

Bảng 18. Tóm tắt kết quả của mô hình đói nghèo tổng thể

Tham số	Ước lượng	Sai số chuẩn	t
Intercept	124.4558	42.68905	2.9154
popdens	0.000597	0.000143	4.1799
lc_natfo	-0.12724	0.023699	-5.3690
lc_arabl	7.034238	2.069287	3.3993
lc_barer	0.071751	0.041173	1.7426
mktparco	-2.67395	0.695687	-3.8436
mktpaym	-7.7E-07	9.8E-07	-0.7855
pctslope1	-0.33163	0.021118	-15.7037
d_dist_m	0.000665	0.000102	6.5066
mainroad	-0.01149	0.00259	-4.4370
minrroad	-0.00324	0.001202	-2.69637
prec_ann	0.000103	0.00114	0.0905
temp_avg	-1.10746	0.410026	-2.7009
sun_annu	-0.00253	0.00222	-1.1412
hum_avg	-0.62011	0.502581	-1.2338

Các câu hỏi chúng tôi muốn quan tâm ở đây là 1) liệu một mô hình vùng có thể tăng khả năng giải thích không và 2) liệu có sự khác nhau về không gian có ý nghĩa giữa các biến giải thích và tỉ lệ đói nghèo hay không? Để kiểm định các câu hỏi này, chúng tôi đã áp dụng một phép

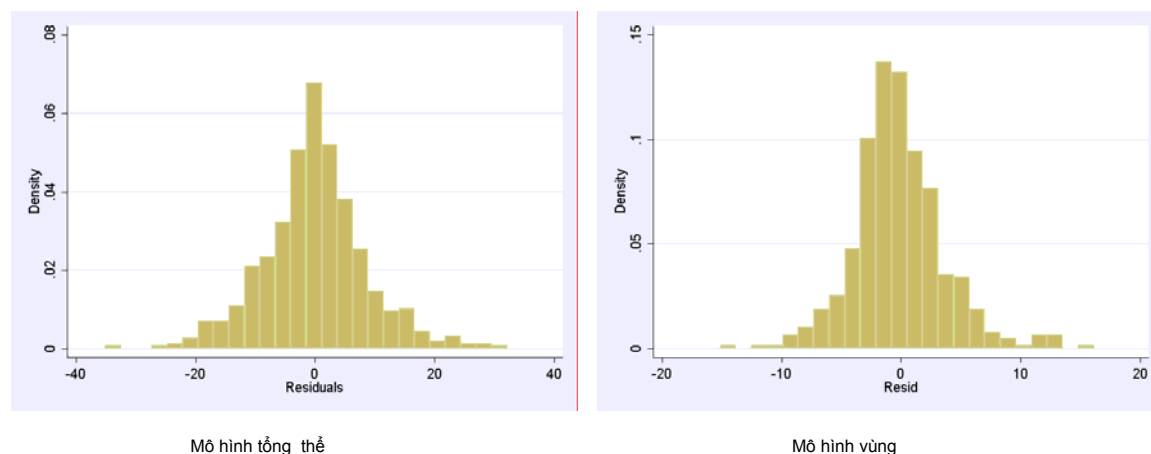
hồi quy quyền số địa lý cho các biến. Bảng 19 là các chỉ số ước lượng chính của mô hình tổng thể so với các kết quả dự đoán của mô hình vùng.

Bảng 19. Tóm tắt kết quả của mô hình thể giới và địa phương

Chỉ số	Kết quả
Global R^2	0.74
Local overall R^2	0.95
Range in local R^2	0.83 – 0.99
Global residual sum of squares	44,908
Local residual sum of squares	8,778

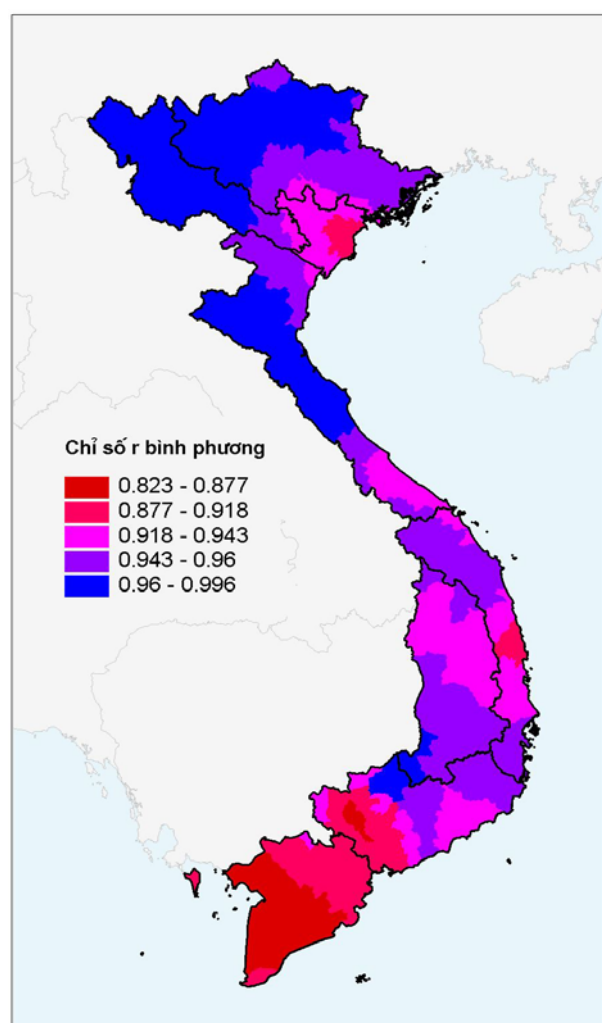
Giá trị chung R^2 của mô hình hồi quy vùng bằng 0,95, cho biết 95% sự biến động về tỉ lệ nghèo đến cấp huyện có thể được giải thích với 14 biến. Điều này cho thấy sự cải thiện đáng kể về khả năng thích hợp của mô hình địa phương khi so sánh với mô hình tổng thể. Kết quả tương tự cũng xuất hiện khi so sánh tổng bình phương phần dư, tổng này nhỏ hơn gần 5 lần trong mô hình vùng so với mô hình tổng thể. Hình 25 biểu diễn sự phân bố của số dư (sự khác biệt giữa tỉ lệ đói nghèo ước lượng và tỉ lệ nghèo thực tế). Cả hai hình đều cho thấy sự phân bố tương đối cân bằng. Nhìn lại kết quả ở trên chúng ta không thấy ngạc nhiên khi thấy mô hình vùng có sự phân bố hẹp hơn so với mô hình tổng thể. Điều này một lần nữa chứng tỏ sự thích hợp và sai số nhỏ hơn trong mô hình hồi quy vùng (local model).

Hình 25. Tóm tắt kết quả của mô hình hồi quy vùng và tổng thể



Tuy nhiên, dường như vẫn có sự biến động đáng kể trong khả năng giải thích của mô hình vùng về mặt không gian. Bản đồ trong Hình 26 cho thấy giá trị R^2 vùng đạt được từ phép phân tích hồi quy vùng. Rõ ràng là, giá trị R^2 đều ở trên giá trị của tổng thể (0,74), trong đó nhìn chung, khu vực miền núi tốt được miêu tả tốt hơn so với miền xuôi và đồng bằng châu thổ. Điều này không quá ngạc nhiên, vì biến sinh thái nông nghiệp và khả năng tiếp cận thị trường thường có ảnh hưởng lớn hơn đến phúc lợi xã hội trong vùng có điều kiện tự nhiên về môi trường khó khăn hơn so với các vùng ít đồi núi hơn, khả năng tiếp cận dễ hơn

Hình 26. Bản đồ phân bố không gian của R^2 từng vùng



Bây giờ, chúng tôi sẽ chuyển sang nghiên cứu sự thay đổi về không gian trong mối quan hệ của các biến.

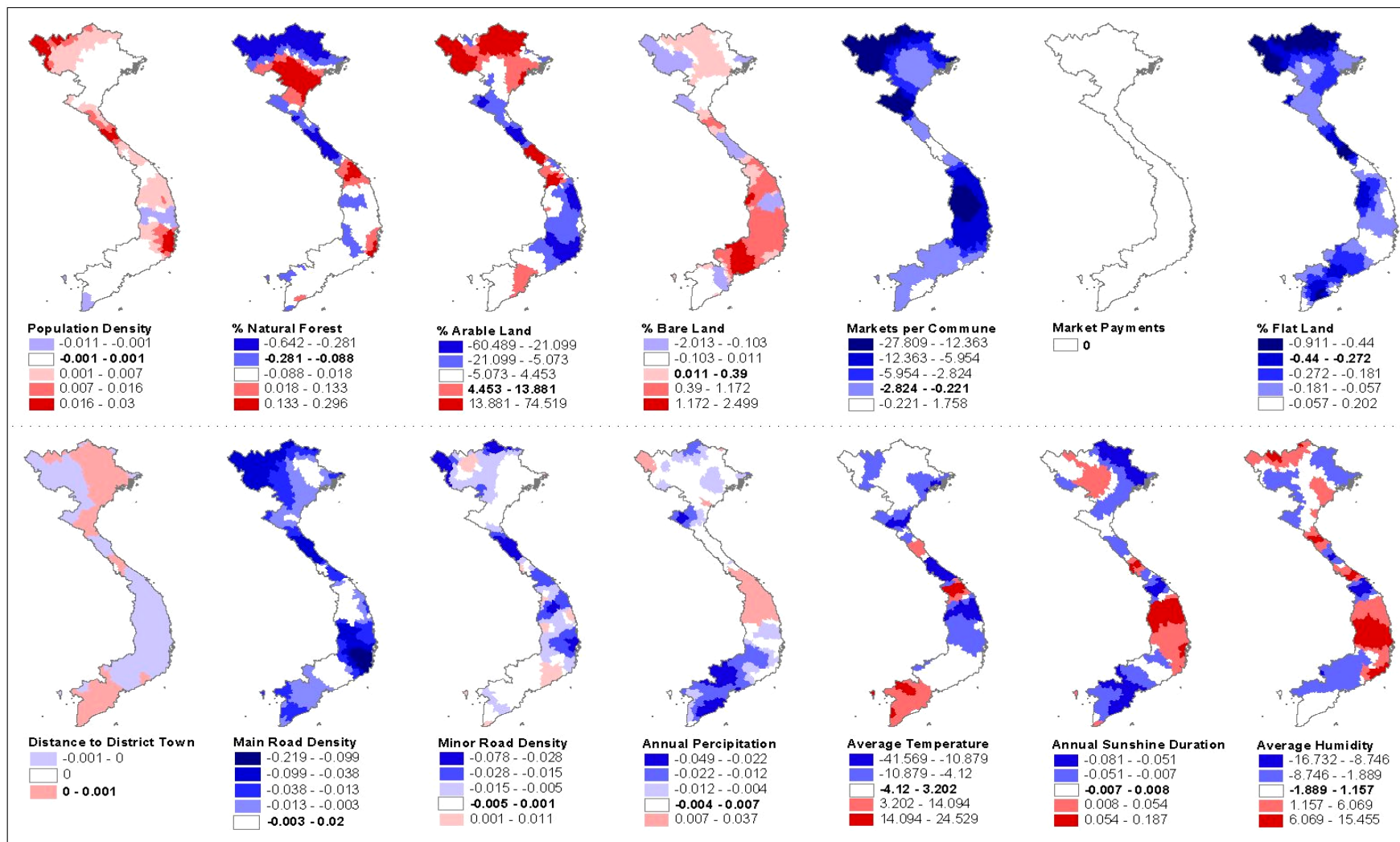
Bảng 20 tổng kết các hệ số vùng được ước lượng từ mô hình hồi quy vùng và Hình 27 biểu diễn sự phân bố không gian của các hệ số trong 14 biến giải thích. Rõ ràng, có thể xác định sự biến động lớn của mối liên hệ trong không gian đối với hầu hết tất cả các tham số, trong đó có nhiều tham số có sự dao động lớn, thường bao gồm sự thay đổi về dấu.

Các phép kiểm định ý nghĩa dựa trên mô phỏng của Monte Carlo được sử dụng để xem xét sự biến đổi không gian có phải do ngẫu nhiên hay là chúng phản ánh đúng sự khác nhau về mặt không gian. Các kiểm định này chứng tỏ rằng tất cả các biến trừ “biến khoảng cách trung bình tới thị trấn huyện” đều có ý nghĩa thống kê trong biến động không gian với độ tin cậy cao (xem Bảng 21).

Bảng 20. Kết quả ước lượng tham số từ mô hình địa phương

Tên biến	Giá trị nhỏ nhất	Giới hạn dưới	Trung vị	Giới hạn trên	Lớn nhất
Intrcent	-1203.02	-53.1318	82.14617	395.9055	1157.351822
pondens	-0.01106	0.000131	0.000232	0.00184	0.030141
lc natfo	-0.64205	-0.08912	-0.031546	0.100038	0.295875
lc arabl	-60.4885	-3.15322	3.700728	9.73332	74.51855
lc barer	-2.01275	-0.05482	0.059338	0.48317	2.499394
mktparco	-27.8087	-7.62955	-2.176135	-1.50498	1.758242
mktpavm	-6.8E-05	-7E-06	-0.000002	0	0.000023
Pctslope1	-0.91052	-0.27857	-0.194937	-0.12296	0.201926
d dist m	-0.00074	0.000122	0.000512	0.000734	0.001523
mainroad	-0.21916	-0.01899	-0.007749	-0.00303	0.020078
minrroad	-0.07833	-0.00879	-0.001636	-0.00053	0.010837
prec ann	-0.04894	-0.01547	-0.004274	0.002622	0.037291
temp avg	-41.5685	-5.20404	-2.874492	1.298637	24.52939
sun annu	-0.08096	-0.0475	-0.011726	0.005215	0.187103
hum_avg	-16.7322	-3.18053	-0.280914	2.020089	15.454768

Hình 27. Bản đồ phân bố không gian các hệ số vùng của các biến độc lập



Bảng 21. Mức ý nghĩa thống kê của biến đổi không gian trong tham số ước lượng

Tham số	Giá trị P	
popdens	0.00	***
lc_natfo	0.00	***
lc_arabl	0.00	***
lc_barer	0.01	***
mktperco	0.00	***
mktpaym	0.03	**
Pctslope1	0.01	***
d_dist_m	0.84	*
mainroad	0.00	***
minrroad	0.00	***
prec_ann	0.00	***
temp_avg	0.00	***
sun_annu	0.00	***
hum_avg	0.00	***

Chú ý: *** có ý nghĩa ở mức 1%

** có ý nghĩa ở mức 5 %, và

* không có ý nghĩa ở mức 10%.

Mặc dù việc giải thích sự biến động trong mối liên hệ trong không gian có thể hơi phức tạp hơn đối với một số biến, song dường như lại dễ dàng hơn đối với các biến khác. Ví dụ, số chợ trung bình trên một xã dường như có tương quan ngược chiều với tỉ lệ nghèo ở các vùng nghèo hơn, vùng cao nguyên xa xôi hơn là ở các vùng đồng bằng. Một kết quả tương tự đối với mật độ đường xá.

Về mối quan hệ của độ che phủ rừng đối với đói nghèo, so sánh Hình 22 với Hình 25 ta thấy diện tích rừng tự nhiên nhìn chung có xu hướng tương quan ngược chiều với đói nghèo ở những vùng miền núi xa xôi do có tỉ lệ che phủ rừng thấp hơn (vùng miền núi phía bắc) so với vùng núi khác (ví dụ ở Tây Nguyên) (xem Hình 24 và Hình 27). Mối tương quan cùng chiều giữa độ che phủ rừng tự nhiên và đói nghèo ở một số vùng đất thấp và các vùng gần trung tâm cho thấy rằng ở các huyện nghèo hơn có tỉ lệ che phủ rừng cao hơn những tỉnh giàu hơn trong cùng một vùng.

Kết quả đáng ngạc nhiên có thể nhận thấy trong các yếu tố không gian của mối liên hệ giữa mật độ dân số và đói nghèo. Trong khi các ý kiến thường cho rằng mật độ dân số tương quan ngược chiều với tỷ lệ đói nghèo (mật độ dân số đông hơn làm tỉ lệ nghèo thấp hơn) nhất là ở các vùng xa xôi hẻo lánh, nhưng bản đồ trong Hình 27 lại cho thấy điều ngược lại. Rõ ràng ở một số vùng xa xôi, mặc dù mật độ dân số tương đối thấp, áp lực dân số vẫn có ảnh hưởng tiêu cực tới phúc lợi mọi người. Điều này có thể giả thích bởi `năng lực thực hiện` thấp của những vùng miền núi, chủ yếu là do năng suất nông nghiệp thấp.

Mặc dù các biến khí hậu cho thấy sự biến động lớn trong mối liên hệ với tỉ lệ đói nghèo, song việc giải thích nó phần nào khó hiểu hơn và cũng không dễ dàng. Tuy nhiên, một vài nhân tố có thể được giải thích một cách dễ hiểu : trong khi về mặt tổng thể thì không có ý nghĩa với đói nghèo nhưng lượng mưa hàng năm ở phần lớn vùng ở Việt Nam ở lại có tương quan ngược với tỉ lệ đói nghèo (lượng mưa cao thì tỉ lệ đói nghèo sẽ thấp), điều này có thể cho biết khu vực duyên hải miền trung không lợi dù có lượng mưa hàng năm cao hơn. Điều này có thể có tương quan chặt chẽ với những vùng thường xuyên bị ảnh hưởng của bão, lụt gây ra những thiệt hại lớn cho vùng.

Yếu tố không gian trong mối liên hệ giữa đói nghèo và các biến kinh tế xã hội, sinh thái nông nghiệp và khả năng tiếp cận thị trường có sự biến động lớn trong không gian. Trong khi một số biến có thể có đóng vai trò quan trọng đối với các điều kiện kinh tế xã hội của một vùng nào đó nhưng đối với vùng khác thì không, hay thậm chí lại có sự tương quan ngược. Các tiếp cận phân tích theo khu vực nhỏ hơn có thể giúp chúng ta thấy những đặc điểm cụ thể hơn của vùng. Những thông tin như vậy có thể rất quan trọng đối với việc lập chính sách hỗ trợ cho người nghèo ở các vùng nhỏ. Trên thực tế, trong một số trường hợp kết luận rút ra từ các kết quả của các mô hình "tổng thể" áp dụng trong các điều kiện "vùng" có thể bị sai lệch. Do đó, những nghiên cứu trong tương lai theo hướng này, ví dụ như phân tích đói nghèo của thành thị và nông thôn theo vùng trong mối liên hệ với từng biến sinh thái nông nghiệp và khả năng tiếp cận thị trường riêng biệt, có thể sẽ cho những kết quả có giá trị hơn về mối liên hệ giữa đói nghèo và môi trường.

6. Tóm tắt và kết luận

Chương này được chia thành ba phần. Phần 6.1 là những tóm tắt sơ lược về đối tượng, phương pháp và kết quả nghiên cứu. Mục đích chính của nghiên cứu là tổng hợp thông tin theo các yếu tố ảnh hưởng đến đói nghèo và bất bình đẳng chứ không phải là nhấn mạnh vào những chính sách cụ thể. Phần 6.2 sẽ đưa ra một số khuyến nghị về mặt chính sách từ các kết quả nghiên cứu. Cuối cùng, Phần 6.3 sẽ đưa ra một số định hướng cho những nghiên cứu trong tương lai có liên quan tới đói nghèo và bất bình đẳng ở Việt Nam.

6.1. Tóm tắt

Giới thiệu chung và phương pháp

Thông tin về sự phân bố về không gian của đói nghèo và bất bình đẳng rất cần thiết bởi nó sẽ giúp những nhà lập chính sách và xây dựng chương trình để xác định những nguyên nhân chính gây ra đói nghèo và nhờ đó đưa ra phương hướng cho các chương trình xóa đói giảm nghèo ở các vùng nghèo nhất. Thông tin về không gian đặc biệt quan trọng hơn ở Việt Nam do có sự khác biệt lớn giữa các vùng và sự cam kết mạnh mẽ của chính phủ trong việc xóa đói giảm nghèo.

Tổng cục Thống kê (GSO) đã thực hiện rất nhiều cuộc điều tra mức sống hộ gia đình nhằm cung cấp thông tin về tỉ lệ đói nghèo đến cấp tỉnh, song các cuộc điều tra này không đưa ra được ước lượng tỉ lệ đói nghèo đến cấp xã, huyện. Bộ Lao động, Thương binh và Xã hội đã đưa ra danh sách các xã nghèo nhất dựa trên số liệu thu thập được từ các cán bộ điều tra thực địa tại xã, song tiêu chí và phương pháp giữa các tỉnh không thống nhất.

Mục đích của nghiên cứu này là : 1) tìm hiểu sự phân bố về không gian của đói nghèo và bất bình đẳng ở Việt Nam sử dụng phương pháp ước lượng diện tích nhỏ (small-area estimation), 2) nghiên cứu các yếu tố địa lý ảnh hưởng đến đói nghèo sử dụng các biến GIS đối với các yếu tố khí hậu nông nghiệp và khả năng tiếp cận thị trường để ước lượng tỉ lệ đói nghèo, và 3) định hướng cho các phương pháp mới, bao gồm phương pháp ước lượng diện tích nhỏ để thu thập thông tin có ích cho các nhà lập chính sách và xây dựng chương trình ở Việt Nam.

Phương pháp ước diện tích nhỏ kết hợp giữa số liệu điều tra mức sống hộ và số liệu Tổng điều tra dân số để ước lượng biến quan tâm (thường là đói nghèo) cho những khu vực nhỏ như các huyện. Đầu tiên, các nhà nghiên cứu sử dụng số liệu của cuộc Điều tra mức sống hộ gia đình Việt Nam 1997-98 đưa ra một phương trình miêu tả mối tương quan giữa mức chi tiêu bình quân đầu người và các đặc điểm của hộ. Sau đó, phương trình này được đưa vào số liệu về các đặc điểm của chính các hộ này trong Tổng Điều tra Dân số và Nhà ở năm 1999 để ước lượng về tỉ lệ đói nghèo đối với mỗi hộ trong Tổng Điều tra dân số. Các kết quả này sẽ

được tổng hợp lại để đưa ra ước lượng tỉ lệ đói nghèo và bất bình đẳng cho cấp xã, huyện và tỉnh.

Bên cạnh đó, nghiên cứu này cũng sẽ phân tích các yếu tố địa lý ảnh hưởng đến của đói nghèo thông qua phân tích hồi quy không gian. Biến phụ thuộc là tỉ lệ đói nghèo cấp tỉnh và các biến giải thích bao gồm các biến đại diện cho địa hình, đất, khí hậu và khả năng tiếp cận thị trường.

Yếu tố không gian ảnh hưởng đến đói nghèo và bất bình đẳng

Trong phần đầu của nghiên cứu này, các nhà nghiên cứu đã sử dụng VLSS để đưa ra một mô hình toán kinh tế về mức chi tiêu bình quân đầu người đối với các hộ gia đình ở nông thôn và thành thị. Như rất nhiều ý kiến, chi tiêu bình quân đầu người có mối liên hệ chặt chẽ với quy mô và thành phần của hộ, trình độ học vấn của chủ hộ và chồng (vợ) của chủ hộ, đặc điểm nhà ở, nguồn nước sinh hoạt, điện sinh hoạt, kiểu nhà vệ sinh và khu vực sinh sống. Song, mô hình này chỉ giải thích được hơn 50% sự biến động về chi tiêu bình quân đầu người đối với nông thôn và thành thị.

Với các yếu tố không gian, tỉ lệ đói nghèo có thể được phân chia như sau :

- Tỉ lệ đói nghèo (P0) cao nhất ở khu vực Đông bắc và Tây bắc dọc biên giới Trung Quốc và CHDCND Lào, ở duyên hải miền Trung và phía Bắc Tây Nguyên .
- Tỉ lệ đói nghèo trung bình ở hai vùng đồng bằng châu thổ chính của Việt Nam là Đồng bằng sông Hồng và Đồng bằng sông Cửu Long.
- Tỉ lệ đói nghèo thấp nhất ở các thành phố lớn, đặc biệt ở Hà Nội, Tp HCM và khu vực miền Đông Nam bộ.
- Tỉ lệ đói nghèo ở thành thị thấp hơn rất nhiều so với ở nông thôn.
- Bản đồ đói nghèo đến cấp xã cho thấy ảnh hưởng của địa hình đồi núi và thậm chí cả đường cao tốc đối với tỉ lệ đói nghèo.
- Có sự khác biệt giữa tỷ lệ đói nghèo cấp huyện trong nghiên cứu này và kết quả của MOLISA

Khoảng tin cậy đối các ước lượng tỷ lệ đói nghèo cho cấp tỉnh và huyện khá hợp lý : một nửa số ước lượng tỷ lệ đói nghèo cấp tỉnh có khoảng tin cậy từ $\pm 4,2$ đến $\pm 5,8\%$ điểm, và một nửa số ước lượng tỷ lệ đói nghèo cấp huyện có khoảng tin cậy từ $\pm 4,4$ đến $\pm 6,9\%$ điểm. Khoảng tin cậy cho các ước lượng về tỷ lệ đói nghèo ở cấp xã cao hơn, cho thấy độ chính xác kém hơn. Một nửa số xã có khoảng tin cậy của tỷ lệ đói nghèo ước lượng từ $\pm 6,6$ đến $\pm 10\%$.

Các nhà nghiên cứu cũng sử dụng hai chỉ số phân tích đói nghèo khác là độ sâu của đói nghèo (P1) và độ trầm trọng của đói nghèo (P2), được ước lượng đến cấp huyện. Hai chỉ

số này có mối liên hệ tương quan rất cao với tỉ lệ đói nghèo (P0), do đó các bản đồ đói nghèo rất giống nhau.

Bản đồ mật độ đói nghèo (số người nghèo trên một đơn vị diện tích) cho thấy mật độ đói nghèo cao nhất ở những nơi tỉ lệ đói nghèo thấp nhất: ở các thành phố và hai đồng bằng châu thổ. Các vùng nghèo nhất, bao gồm Đông bắc, Tây Bắc và Tây Nguyên nơi dân cư thưa thớt thì lại có mật độ đói nghèo cao hơn nhiều. Ngược lại, những vùng đô thị đông đúc và đồng bằng thì lại có số người nghèo nhiều nhất cho dù tỷ lệ đói nghèo của những vùng này thấp.

Nghiên cứu này cũng đưa ra ba bất bình đẳng trong chi tiêu bình quân đầu người là hệ số Gini, chỉ số Theil L và chỉ số Theil T. Chúng ta có thể tóm tắt kết quả như sau :

- Ba chỉ số bất bình đẳng có tương quan chặt chẽ với giá trị $R^2 = 0,97 - 0,98$.
- Tỉ lệ bất bình đẳng lớn nhất ở các thành phố lớn và một số vùng thuộc Đông bắc, Tây bắc và Tây Nguyên.
- Tỉ lệ bất bình đẳng thấp nhất ở vùng Đồng bằng sông Hồng, tiếp đến là vùng Đồng bằng sông Cửu Long.
- Hơn 3/4 tỉ lệ bất bình đẳng là do sự khác nhau trong tỉnh hơn là giữa các tỉnh.
- Khoảng 2/3 bất bình đẳng là do sự khác nhau trong huyện hơn là giữa các huyện

Đánh giá mối quan hệ giữa nghèo đói, bất bình đẳng, chi tiêu bình quân đầu người và tỉ lệ đô thị hoá đến cấp huyện, kết quả nghiên cứu cho thấy :

- Tỉ lệ đói nghèo (P0) huyện và mức chi tiêu bình quân đầu người trung bình huyện có mối liên hệ khá chặt chẽ, với mức chi tiêu bình quân đầu người giải thích được 96% sự biến động bất bình đẳng giữa các huyện.
- Nhìn chung, mức chi tiêu bình quân đầu người cao tương ứng với tỉ lệ bất bình đẳng cao, song ở một số huyện nghèo hơn sẽ có cả một số huyện có tỉ lệ bất bình đẳng rất cao.
- Tỉ lệ bất bình đẳng lớn nhất ở các huyện có tỉ lệ nghèo đói rất thấp hoặc rất cao, mặc dù sự tương quan không mật thiết.
- Khi dân số thành thị tăng lên, tỉ lệ nghèo đói giảm
- Khi dân số thành thị tăng lên, tỉ lệ bất bình đẳng cũng tăng lên tới một điểm, sau điểm đó mức độ đô thị hoá tăng tương ứng với tỉ lệ bất bình đẳng giảm. Nói cách khác, các huyện có tỉ lệ bất bình đẳng cao nhất là các tỉnh vừa có dân thành thị và nông thôn.

Các yếu tố địa lý ảnh hưởng đến đói nghèo

Phần này phân tích các yếu tố địa lý tác động đến đói nghèo sử dụng phương pháp phân tích hồi quy không gian. Biến phụ thuộc là tỉ lệ đói nghèo đến cấp huyện (PO) và các biến giải thích là biến GIS như độ cao so với mặt nước biển, độ dốc, loại đất, đất mặt, lượng mưa, số ngày nắng, khoảng cách tới thị xã và thành phố. Các mô hình khác nhau được sử dụng để ước lượng tỉ lệ đói nghèo cho thành thị và nông thôn.

Các kiểm định thống kê trên mô hình nông thôn cho thấy có sự tương quan của các sai số. Điều này có nghĩa là một số biến địa lý không có trong mô hình tạo ra sai số ở các huyện lân cận có tương quan với nhau. Chúng tôi đã giảm bớt hiện tượng này bằng cách áp dụng phương pháp bình phương nhỏ nhất tổng hợp (generalized least squares) với một ma trận quyền số không gian (spatial weighting matrix). Một mô hình đầy đủ (inclusive model) với 32 biến giải thích đã giải thích 80% sự biến động tỉ lệ đói nghèo cấp huyện. Các yếu tố sau đây cũng có tương quan cùng chiều tới nghèo đói ở nông thôn : đất trồng và đất đá, địa hình dốc đứng, loại đất mặt nhiễm axit, đất cát, đất muối và khoảng cách tới thị trấn ít nhất 10.000 người. Độ cao so với mặt nước biển, lượng mưa trung bình hàng năm và số giờ nắng hàng năm không có ảnh hưởng đáng kể về mặt thống kê. Sau khi bỏ qua tất cả các yếu tố và các biến không có ý nghĩa, chúng tôi thu được một mô hình chọn lọc hơn về tình trạng đói nghèo ở nông thôn, trong đó 12 biến giải thích đã giải thích 74% sự biến động tỉ lệ đói nghèo đến cấp huyện. Trong mô hình này, khoảng cách tới thành phố ít nhất 100.000 người cũng có ý nghĩa.

Các kiểm định về sự tương quan không gian trong mô hình đói nghèo cho thành thị không rõ ràng, song điều này lại gợi ý cho chúng tôi nên sử dụng lại mô hình sai số không gian. Mô hình đầy đủ với 32 biến giải thích chỉ giải thích được 38% sự biến động về tỉ lệ đói nghèo ở thành thị. Điều này có nghĩa là tỉ lệ đói nghèo ở thành thị ít bị ảnh hưởng bởi các điều kiện khí hậu nông nghiệp và khả năng tiếp cận thị trường hơn so với nông thôn. Tỉ lệ đói nghèo thành thị chịu ảnh hưởng bởi các yếu tố như đất có thể canh tác, khoảng cách tới thị xã ít nhất 10.000 người, khoảng cách tới thành phố ít nhất 100.000 người, song không bị ảnh hưởng bởi địa hình đất đỏ bazan. Trong mô hình chọn lọc này, sáu biến chỉ giải thích 29% sự biến động về tỉ lệ đói nghèo ở thành thị.

Biến động về mặt không gian của các yếu tố ảnh hưởng đến đói nghèo.

Ảnh hưởng của các biến khí hậu nông nghiệp và tiếp cận thị trường tới nghèo đói có sự khác nhau giữa các vùng. Bằng cách áp dụng phép phân tích hồi quy vùng có sử dụng quyền số không gian, chúng tôi nghiên cứu ảnh hưởng của sự biến động về mặt không gian trong mối liên hệ của 14 biến giải thích tới tỉ lệ đói nghèo.

Các kết quả phân tích này chứng tỏ hai điều: thứ nhất, có khác nhau đáng kể trong cách các biến giải thích tác động đến đói nghèo. Trong khi sự liên quan của yếu tố tiếp cận thị trường và tiếp cận các nguồn tài nguyên như rừng tự nhiên đối với tỉ lệ nghèo đói ở mức cao nhất ở

các vùng khó khăn, sự biến động trong tương quan của khí hậu đối với nghèo đói phản ánh sự tổn thương do thiên tai như lũ lụt hay bão. Thứ hai, một mô hình cho phép có sự biến động lớn về mặt không gian trong các mối liên hệ sẽ mô tả tốt hơn mối quan hệ phức tạp giữa nghèo đói và môi trường. Các kết quả phân tích cho thấy những chỉ số của mô hình vùng (local model) về sự phù hợp của mô hình (goodness of fit) ở đâu cũng cao hơn trong mô hình tổng thể (global model), nhưng vẫn có sự khác nhau rất lớn giữa các vùng về vấn đề là làm thế nào mô hình có thể tái tạo dữ liệu. Nhìn chung, các vùng có địa hình khó khăn hơn sẽ có mức độ phù hợp tốt hơn trong mô hình, cho thấy các yếu tố sinh thái nông nghiệp và khả năng tiếp cận thị trường có ảnh hưởng rất lớn đến phúc lợi xã hội ở những vùng miền núi so với những vùng có điều kiện môi trường thuận lợi hơn.

6.2. Kết luận

Các kết quả ước lượng về tỷ lệ đói nghèo cấp huyện trong nghiên cứu này khác với các kết quả của MOLISA. Mặc dù MOLISA sử dụng chỉ số phúc lợi (welfare), đường chuẩn nghèo khác, phương pháp điều chỉnh chỉ tiêu giữa các vùng khác và đo lường đói nghèo cấp hộ chứ không phải cá nhân, nhưng có điều đáng ngạc nhiên là chỉ có một tương quan yếu về tỷ lệ đói nghèo ước lượng từ hai phương pháp.

Chênh lệch lớn giữa tỉ lệ đói nghèo giữa các huyện và xã . Một trong những khía cạnh mà bản đồ đói nghèo khai thác là sự biến động lớn giữa tỉ lệ đói nghèo ở các vùng khác nhau trên cả nước. Tại một số vùng, đặc biệt là các vùng miền núi xa xôi, tỉ lệ dân sống dưới mức nghèo đói là hơn 90%. Tại các vùng khác, nhất là các vùng trong hoặc gần các trung tâm đô thị lớn, chỉ có dưới 5% dân số là người nghèo.

Mặc dù tỉ lệ đói nghèo giữa các vùng trong cả nước có sự khác nhau lớn, nhưng mức độ bất bình đẳng lại tương đối thấp. Một biến có thể ước lượng một nước có sự biến động lớn về tỉ lệ đói nghèo sẽ có tỉ lệ bất bình đẳng cao, song tỉ lệ bất bình đẳng ở Việt Nam vẫn tương đối thấp so với tiêu chuẩn quốc tế. Nguyên nhân là do các vùng nghèo nhất thường lại thưa thớt dân cư, do đó sẽ không có tác động lớn đến tỉ lệ bất bình đẳng của quốc gia. Phần lớn dân số nông thôn ở Việt Nam tập trung ở Đồng bằng sông Cửu Long và Đồng bằng sông Hồng, khu vực có tỉ lệ bất bình đẳng thấp. Điều đáng ngạc nhiên hơn nữa là thực tế rằng phần lớn tỉ lệ bất bình đẳng lại tập trung ở các huyện riêng rẽ. Sự biến động trong chỉ tiêu bình quân đầu người giữa các vùng chỉ chiếm 1/3 tổng tỉ lệ bất bình đẳng trong cả nước.

Sự bất bình đẳng cao không chỉ tồn tại ở thành thị và các vùng sản xuất hàng hóa. Người ta thường cho rằng ở Việt Nam (cũng giống như các nước đang phát triển khác), tỉ lệ bất bình đẳng tập trung chủ yếu ở các vùng thành thị và các vùng nông thôn nổi bật với mô hình nông nghiệp thương mại quy mô lớn. Cơ sở là, tỉ lệ bất bình đẳng thể hiện ở mức tăng trưởng kinh tế theo sản phẩm, với một số hộ gia đình đã tận dụng tốt cơ chế thị trường và có thu nhập cao hơn nhiều so với mức trung bình. Kết quả nghiên cứu của chúng tôi cho thấy tỉ lệ bất bình

đẳng cao hơn ở các vùng thành thị, song chúng tôi cũng nhận thấy rằng tỉ lệ này cũng khá cao ở các vùng nông thôn, thậm chí ở các vùng dân cư thưa thớt và thu nhập một phần từ các hoạt động nông nghiệp.

Sự khác biệt về tỉ lệ đói nghèo giữa các huyện chủ yếu là do sự chênh lệch về mức chi tiêu bình quân đầu người trung bình, không phải do sự bất bình đẳng. 96% biến động về tỉ lệ đói nghèo cấp huyện có thể được giải thích bằng sự chênh lệch trong mức chi tiêu bình quân đầu người, với sự khác biệt về tỉ lệ bất bình đẳng chỉ chiếm dưới 3%. Nguyên nhân là do tỉ lệ bất bình đẳng ở cấp tỉnh huyện tương đối thấp và ổn định. Tuy nhiên, chúng ta hoàn toàn có thể giảm tỉ lệ đói nghèo cấp tỉnh bằng cách giảm tỉ lệ bất bình đẳng, song để phân biệt rõ đâu là vùng có tỉ lệ đói nghèo cao và đâu là vùng có tỉ lệ đói nghèo thấp lại không phải dễ dàng. Nếu chỉ tiêu từng vùng phản ánh sự thay đổi xảy ra trong suốt thời gian đó, khi đó, tỉ lệ đói nghèo giảm phần lớn là do tăng trưởng kinh tế hơn là nhờ sự cải thiện về phân bổ thu nhập. Kết quả này nhấn mạnh đến tầm quan trọng của các chính sách và các chương trình thúc đẩy tăng trưởng thu nhập hộ gia đình như một chiến lược xóa đói giảm nghèo.

Phần lớn người nghèo sống ở các vùng không nghèo. Trong khi mật độ người nghèo thấp nhất ở những vùng có tỉ lệ đói nghèo cao nhất (các vùng nông thôn miền núi), thì mật độ đói nghèo lại cao nhất ở những vùng có tỉ lệ đói nghèo thấp (các đô thị và đồng bằng nông thôn). Số người nghèo trong những vùng có tỷ lệ đói nghèo cao là tương đối thấp vì mật độ dân số ở những vùng này cũng thấp. Ngược lại, hầu hết các hộ nghèo nông thôn tập trung ở Đồng Bằng sông Hồng và Đồng Bằng sông Cửu Long. Mặc dù những vùng này có tỷ lệ đói nghèo tương đối so với các vùng nông thôn khác, nhưng mật độ dân số cao nên hầu hết người nghèo sống ở hai vùng này.

Phần lớn sự biến động về tỷ lệ đói nghèo nông thôn có thể được giải thích bởi các biến nông nghiệp, khí hậu và tiếp cận thị trường. Không ngạc nhiên khi các biến khí hậu, nông nghiệp có thể giải thích được sự biến động tỷ lệ đói nghèo ở nông thôn nhưng thật không ngờ các biến này có thể giải thích $\frac{3}{4}$ sự biến động của tỷ lệ đói nghèo nông thôn. Trong khi đó, chưa đến $\frac{2}{5}$ sự biến động của tỷ lệ đói nghèo thành thị được giải thích bởi các biến này.

6.3 Khuyến nghị về mặt chính sách và cho các chương trình

Mục đích chính của nghiên cứu này là nghiên cứu về đói nghèo và bất bình đẳng giữa các khu vực, với ý tưởng là những thông tin này sẽ có ích cho các chương trình xóa đói giảm nghèo. Nghiên cứu không được thiết kế để đưa ra những lựa chọn chính sách cụ thể nhằm xóa đói giảm nghèo. Tuy nhiên, các kết quả cung cấp một số các khuyến nghị gián tiếp về mặt chính sách và cho các chương trình. Trong phần này chúng ta sẽ thảo luận một số các khuyến nghị này.

Người nghèo ở đâu?

Một ứng dụng rõ ràng của các kết quả trình bày trong nghiên cứu này là cung cấp thêm thông tin về sự phân bố đói nghèo của các khu vực xét về mặt không gian cho các chương trình xóa

đói giảm nghèo. Những kết quả không chỉ cung cấp những thông tin về sự phân bố đói nghèo của Việt Nam mà còn cho biết mức độ chính xác của các thông tin này. Hơn nữa, thông qua việc xây dựng các chỉ số đói nghèo khác nhau, báo cáo cho phép những người xây dựng chương trình đưa ra những sự hỗ trợ cho những huyện có khoảng cách nghèo đói và mức độ trầm trọng của nghèo đói cao nhất.

Hỗ trợ cho những vùng nghèo hay người nghèo?

Nếu hầu hết người nghèo sống ở những vùng có tỷ lệ đói nghèo thấp, thì những khuyến nghị gì cho chương trình xoá đói giảm nghèo mục tiêu? Đặc biệt, các chương trình xoá đói giảm nghèo có nên tập trung vào những vùng có mật độ đói nghèo cao nhất? Câu trả lời phụ thuộc vào đặc điểm của các chương trình, và chúng ta sẽ thảo luận dưới đây.

Một số chương trình không hoàn toàn nhằm mục đích nâng cao thu nhập của tất cả các hộ gia đình trong một vùng. Ví dụ như đường xá được nâng cấp, sức khoẻ người dân được chăm sóc tốt hơn và hỗ trợ tài chính cho chính quyền địa phương. Giả sử chương trình này được ấn định một mức trợ cấp nhất định cho mỗi người dân, chương trình này sẽ có tác động lớn tới đói nghèo nếu tập trung vào các vùng nghèo. Ở các vùng này, tỉ lệ người nghèo cao hơn do đó tỉ lệ người nghèo hưởng lợi từ các chương trình cao hơn. Bằng cách này, chính phủ sẽ đạt được tỉ lệ giảm đói nghèo tính bình quân trên một đôla đầu tư cao hơn. Điều này hoàn toàn đúng trong trường hợp nếu mục tiêu là giảm tỉ lệ đói nghèo P1 (độ sâu của đói nghèo) và cũng có thể đúng nếu mục tiêu là giảm P0 (tỉ lệ đói nghèo).

Các chương trình khác lại hướng vào các hộ nghèo (như trợ cấp, cung cấp lương thực cho việc làm hay miễn phí các dịch vụ xã hội). Nếu mục tiêu là cung cấp mức hỗ trợ tương tự đối với mỗi người dân nghèo, chương trình nên triển khai cho những vùng có nhiều người nghèo, mặc dù chi phí bình quân mỗi người ở những vùng có tỉ lệ đói nghèo cao sẽ tăng lên.

Tất nhiên, những định hướng trên đều giả định rằng chi phí bình quân đầu người của các chương trình là cố định, nói cách khác là chi phí không bị tác động bởi mật độ dân số. Một vài chương trình, như là chương trình điện khí hoá và khuyến nông, sẽ cần chi phí bình quân đầu người cao hơn ở những vùng có mật độ người nghèo thấp. Các chương trình khác, đặc biệt là các chương trình điện, đường, trường, trạm, có thể sẽ có chi phí cao hơn ở một vùng có mật độ người nghèo cao.

Có phải do địa lý nên không thể tạo ra sự phát triển ở vùng núi?

Phân tích các yếu tố địa lý tác động đến đói nghèo cho thấy $\frac{3}{4}$ biến động của đói nghèo nông thôn ở các huyện có thể giải thích bằng một số các biến về khí hậu nông nghiệp và khả năng tiếp cận thị trường. Kết quả này phần nào cho thấy sự khó khăn của các chương trình do chúng ta không thể thực hiện chính sách can thiệp tác động trực tiếp tới các biến khí hậu nông nghiệp. Do đó, những người sống ở các vùng có độ dốc lớn và đất đai cằn cỗi có thể bị rơi vào vòng luẩn quẩn của nghèo đói, rất khó có thể thoát ra được.

Khả năng tiếp cận thị trường là một trong các biến địa lý có thể tác động bằng các chính sách. Mặc dù chính phủ không thể thu hẹp khoảng cách thực tế của các vùng xa xôi tới các thành phố, song có thể giảm bớt thời gian đi lại và chi phí. Đường xá ngăn lại sẽ làm giảm chi phí của hàng hoá sản xuất ở các vùng khác (như mặt hàng gạo) tiêu thụ ở những vùng nghèo và cạnh tranh với các hàng địa phương. Tuy nhiên, điều này sẽ được bù đắp bằng việc nâng cao khả năng tiếp cận các thị trường thu nhập cao ở bên ngoài.

Hơn nữa, yếu tố địa lý chỉ là một nhân tố cản trở xoá đói giảm nghèo khi mọi người không thể di cư. Khi những người di cư có thể tăng mức sống của họ mà không ảnh hưởng xấu tới những người khác, việc di cư có thể là một công cụ hiệu quả để giảm đói nghèo. Ý nghĩa của vấn đề này là Chính phủ nên coi sự di cư như một chiến lược phát triển, đặc biệt đối với những huyện có điều kiện sinh thái nông nghiệp bất lợi. Việc cho phép sự di cư sẽ cho phép những người dân ở các vùng có điều kiện sinh thái nông nghiệp bất lợi có điều kiện tăng thu nhập và giảm nghèo. Mặc dù, những ban đầu người di cư từ nông thôn ra thành phố sẽ nghèo hơn những người xung quanh, do đó sẽ làm cho số người nghèo ở thành thị tăng lên, vấn đề ở đây là liệu mức sống của những người di cư có tốt hơn không.

Cuối cùng, một điều rất quan trọng là không nên coi địa lý là một yếu tố cản trở bất kỳ sự phát triển nào ở những vùng có điều kiện tự nhiên bất lợi. Những nghiên cứu khác cũng chỉ ra rằng thậm chí tăng trưởng kinh tế và giảm đói nghèo vẫn có thể xảy ra ở những vùng có điều kiện tự nhiên khó khăn như miền núi phía Bắc. (Nhóm làm việc về đói nghèo, 1999). Thực tế là các yếu tố về khí hậu, nông nghiệp có ảnh hưởng tới đói nghèo giữa các huyện tại một điểm trong một thời gian nào đó không có nghĩa là nó có ảnh hưởng liên tục tới đói nghèo đối với một huyện nào đó.

Tăng trưởng và bình đẳng

Tại Việt Nam cũng như ở nơi khác, luôn có tranh luận giữa những người ủng hộ các chính sách và chương trình xoá đói giảm nghèo thông qua hỗ trợ trực tiếp cho người nghèo và những người ủng hộ những chính sách và chương trình tăng trưởng kinh tế như là một chiến lược xoá đói giảm nghèo. Nghiên cứu này cho thấy hầu hết (96%) của các biến động tỷ lệ đói nghèo giữa các huyện có thể giải thích bởi sự khác biệt chỉ tiêu bình quân đầu người trung bình cấp huyện. Tất nhiên, đói nghèo chúng ta có thể giảm đói nghèo bằng cách giảm sự bất bình đẳng, nhưng trong thực tế đây không phải là tiêu chuẩn để phân biệt huyện có tỷ lệ đói nghèo thấp và huyện có tỷ lệ đói nghèo cao ở Việt Nam. Nếu xu hướng giữa các vùng phản ánh sự thay đổi theo thời gian, vì thế khuyến nghị ở đây là xoá đói giảm nghèo có thể giảm nhanh hơn từ kết quả của tăng trưởng kinh tế theo cách tiếp cận rộng chứ không chỉ tập trung vào việc phân bổ thu nhập. Hơn nữa, điều này cũng phù hợp với kết quả khi chúng ta so sánh của cuộc điều tra Mức sống dân cư Việt Nam năm 1993 và năm 1998 (xem Nhóm công tác đói nghèo, 1999). Kết quả này cho thấy tầm quan trọng của các chính sách và chương trình tăng thu nhập của hộ như là một chiến lược xoá đói giảm nghèo.

6.4. Định hướng cho nghiên cứu trong tương lai

Đói nghèo ở cấp hộ gia đình có thể được giải thích một cách hợp lý dựa trên các đặc điểm cơ bản của hộ. Hơn 50% biến động trong chỉ tiêu bình quân đầu người có thể giải thích qua phân tích 17 đặc điểm của hộ trong phiếu phỏng vấn Tổng điều tra dân số. Một phiếu phỏng vấn được thiết kế tập trung những đặc điểm phân biệt hộ nghèo và không nghèo thậm chí có thể dự đoán mức chỉ tiêu tốt hơn. Điều này cho thấy triển vọng trong việc tổ chức một cuộc điều tra (hoặc các chỉ tiêu trong điều tra) tập trung vào các đặc điểm của hộ phù hợp với chỉ tiêu hoặc thu nhập. Kết quả này có thể được sử dụng trong quản lý đói nghèo, đánh giá dự án hoặc lập mục tiêu cấp hộ. Ở Việt Nam đã có một vài công trình nghiên cứu được tiến hành với chủ đề này (xem Minot and Baulch, 2002), song vẫn cần có những nghiên cứu sâu rộng hơn để tìm ra các đặc điểm nổi bật nhất và đánh giá mục tiêu của nghiên cứu.

Cần phải có những nghiên cứu nữa để đánh giá sự khác nhau giữa các kết quả về đói nghèo trong nghiên cứu của chúng tôi với của MOLISA. Những ước lượng trong nghiên cứu này có thể chưa hoàn toàn chính xác nếu, ở một số vùng, đói nghèo không dự đoán tốt trên cơ sở đặc điểm của hộ. Tuy nhiên, ước lượng của MOLISA cũng có thể có những sai lầm do sự không thống nhất về tiêu chí xác định đói nghèo và cách thức giám sát đói nghèo (Conway, 2001). Một cách tiếp cận có thể nghiên cứu sâu hơn số huyện trong đó hai phương pháp cho những kết quả rất khác nhau.

Ước lượng diện tích nhỏ là một công cụ hiệu quả để phân tích sự phân bố nghèo đói và bất bình đẳng về mặt không gian. Kết quả trong báo cáo này cho thấy tiềm năng đáng kể trong việc sử dụng phương pháp ước lượng theo diện tích nhỏ và số liệu điều tra dân số để có thể hiểu rõ hơn về sự phân bố về mặt không gian của đói nghèo và bất bình đẳng. Số liệu Tổng điều tra dân số cung cấp số liệu ở khu vực nhỏ hơn, điều này sẽ ngày càng trở nên cần thiết cho việc phân tích chính sách và sự phân quyền cho những vùng xa xôi.

Tuy nhiên, phương pháp ước lượng diện tích nhỏ không thể dễ dàng cập nhật bản đồ nghèo đói. Mặc dù phương pháp ước lượng diện tích nhỏ rất có giá trị trong việc xây dựng các bản đồ đói nghèo và các thông tin khác về sự phân bố không gian của đói nghèo và bất bình đẳng, song nó không mấy hiệu quả khi tổng hợp đánh giá đói nghèo ở cấp xã và huyện trong phạm vi cả nước hàng năm. Nếu phân tích này sử dụng số liệu điều tra dân số trong giai đoạn hai, nó có thể được cập nhật 10 năm một lần. Số liệu từ điều tra dân số nông nghiệp có thể được sử dụng để cập nhật các dự đoán đói nghèo ở nông thôn 5 năm một lần. Tuy nhiên, các cuộc điều tra hộ hàng năm, giống như các cuộc điều tra của Tổng cục Thống kê, sẽ chỉ có thể cập nhật các phương trình dự đoán chứ không phải các ước lượng về đói nghèo. Để cập nhật ước lượng tỉ lệ đói nghèo cấp huyện chỉ cần một cuộc điều tra đơn giản (với một mẫu Phiếu phỏng vấn tương tự như phiếu điều tra dân số), song với một mẫu lớn có thể tới 600.000 hộ.

Phương pháp ước lượng diện tích nhỏ cũng có thể được áp dụng trong nghiên cứu sự phân bố không gian về dinh dưỡng, thương mại nông nghiệp hoặc bất cứ biến nào khác có thể được ước lượng dựa trên các đặc điểm của hộ. Mặc dù nghiên cứu này sử dụng phương pháp

ước lượng diện tích nhỏ để nghiên cứu yếu tố không gian tác động đến tỉ lệ đói nghèo và dinh dưỡng, nhưng phương pháp này còn có thể được sử dụng để khai thác yếu tố không gian đối với các biến quan tâm khác. Ví dụ, nếu tình trạng thiếu calo hoặc thiếu vi dinh dưỡng có thể được ước lượng thông qua việc phân tích các đặc điểm của hộ trong điều tra dinh dưỡng, kết quả có thể được áp dụng vào số liệu điều tra dân số để tổng hợp thông tin chi tiết về sự phân bố về mặt không gian các vấn đề đó. Tương tự như vậy, các biến khác như mức độ đa dạng hoá thu nhập, khả năng tổn thương đối với những vùng bị thiên tai hoặc sự tham gia các hoạt động sản xuất nông nghiệp thương mại có thể được lập bản đồ theo cách tương tự nếu chúng có thể được ước lượng với độ chính xác tương đối từ các đặc điểm của hộ trong số liệu điều tra dân số.

TÀI LIỆU THAM KHẢO

- Anselin, L., 1988. *Toán kinh tế không gian : Phương pháp và mô hình (Spatial Econometrics: Methods and Models)*. Dordrecht: Kluwer.
- Baker, J. and Grosh, M., 1994, “Xoá đói giảm nghèo thông qua tiếp cận yếu tố địa lý: kết quả ra sao?” (“Poverty reduction through geographic targeting: how well does it work?”) *Phát triển thế giới (World Development)*, Vol. 22, No. 7: 983-995
- Baulch, B., 2002 “Mục tiêu và giám sát đói nghèo bằng cách sử dụng đường cong ROC: Trường hợp của Việt Nam”, Bài viết 161, Viện Nghiên cứu Phát triển
- Baulch, B., Truong, C., Haughton, D. and Haughton, J., 2003. “Sự phát triển của các dân tộc thiểu số ở Việt Nam: một cách nhìn kinh tế xã hội” (“Ethnic minority development in Vietnam: a socio-economic perspective.”)
- Bigman, D. and Fofack, H., 2000, *Các yếu tố địa lý trong Xoá đói giảm nghèo : phương pháp và áp dụng (Geographic Targeting for Poverty Alleviation: Methodology and Applications)*, Washington DC: World Bank Regional and Sectoral Studies
- Brunsdon C., Fotheringham A. S. and Charlton M. E., 1996. “Hồi quy quyền số địa lý : Một phương pháp không dừng” (“Geographically Weighted Regression: A method for exploring spatial nonstationarity”), *Geographic Analysis* Vol. 28, No. 4: 281 – 298
- Conway. T., 2001 “Sử dụng số liệu của Chính Phủ để đặt ra các hoạt động cho xã nghèo và giám sát công tác xoá đói giảm nghèo: Xem xét lựa chọn cho dự án Phát triển ở Bắc Cạn và Cao Bằng” (“Using government data to target activities to poor communes and monitor poverty reduction: a review of options for the Cao Bang-Bac Kan Rural Development Project”), Hà nội: Ủy ban cộng đồng Châu âu
- Elbers, C., Lanjouw, J. and Lanjouw, P., 2003, “Ước lượng vi mô về đói nghèo và bất bình đẳng” (“Micro-level estimation of poverty and inequality.”) *Econometrica* 71 (1): 355-364.
- Epprecht M. and Müller D., 2003, “Liên kết con người và không gian : GIS và kỹ thuật phân tích không gian đối với các nhà kinh tế nông nghiệp” (Linking People and the Landscape: GIS and Spatial Analytical Techniques for Agricultural Economists), bài phát biểu tại Hội nghị quốc tế Hiệp hội các nhà kinh tế nông nghiệp quốc tế, Durban, 2003

Fotheringham A. S., Brunsdon M. và Charlton M., 2002. *Hồi quy quyền số không gian : phân tích mối liên hệ biến đổi trong không gian (Geographically Weighted Regression: the analysis of spatially varying relationships)*, Chichester: Wiley.

Tổng cục thống kê. 2000. *Điều tra mức sống hộ gia đình Việt Nam 1997-1998 (Vietnam Living Standards Survey 1997-1998)*. Nhà xuất bản thống kê, Hà nội.

Henninger, N. và M. Snel. 2002. *Người nghèo ở đâu ? Kinh nghiệm trong phát triển và sử dụng bản đồ đói nghèo. (Where are the poor? Experiences with the development and use of poverty maps)*. Viện tài nguyên thế giới, Washington, D.C. và UNEP-GRID/Arendal, Arendal, Na uy.

Hentschel, J., Lanjouw, J., Lanjouw, P. và Poggi, J., 2000, “Kết hợp điều tra dân số và số liệu điều tra để xác định phạm vi không gian của đói nghèo : nghiên cứu dựa trên đối tượng của Ecuador” (“Combining census and survey data to trace the spatial dimensions of poverty: a case study of Ecuador”), *World Bank Economic Review*, Vol. 14, No. 1: 147-65

Heywood I., Cornelius S., và Carver S., 2002, *Giới thiệu hệ thống thông tin địa lý, tái bản lần thứ hai (An Introduction to Geographical Information Systems, Second Edition)*, Prentice Hall, Harlow.

Kanbur, R., 2002, “Tầm quan trọng chính sách của việc phân tích sự bất bình đẳng” (“Notes of the Policy Significance of Inequality Decompositions”) Mimeo, Trường đại học Cornell

Minot, N., 1998, “Tổng hợp bản đồ đói nghèo : áp dụng tại Việt Nam” (“Generating disaggregated poverty maps: An application to Viet Nam”). Markets and Structural Studies Division, Discussion Paper No. 25.. International Food Policy Research Institute, Washington, D.C.

Minot, N., 2000, “Xây dựng bản đồ đói nghèo: áp dụng tại Việt Nam” (“Generating disaggregated poverty maps: an application to Vietnam”), *World Development*, Vol. 28, No. 2: 319-331

Minot, N. và B. Baulch. 2002a. “Phân bố không gian của đói nghèo ở Việt Nam và triển vọng xoá đói giảm nghèo” (“The spatial distribution of poverty in Vietnam and the potential for targeting.”) Discussion Paper No. 42. Markets and Structural Studies Division, International Food Policy Research Institute, Washington, D.C.

Minot, N. và B. Baulch. 2002b. “Bản đồ đói nghèo với sự kết hợp các số liệu điều tra dân số” (“Poverty mapping with aggregate census data: What is the loss in precision.”) Discussion Paper No. 49. Markets and Structural Studies Division, International Food Policy Research Institute, Washington, D.C.

Tổ chức hành động chống đói nghèo (Poverty Working Group), 2000, Báo cáo phát triển Việt Nam : Tấn công đói nghèo (*Vietnam Development Report: Attacking Poverty*), Báo cáo kết

hợp giữa Tổ chức phi chính phủ hành động chống đói nghèo do chính phủ tài trợ tại cuộc họp các nhóm tư vấn cho Việt Nam (A Joint Report of the Government of Vietnam-Donor-NGO Poverty Working Group presented to the Consultative Group Meeting for Vietnam)

Ravallion, M. , “So sánh đói nghèo” (“Poverty Comparisons”), *Living Standard Measurement Working Paper* No. 88, Washington DC: World Bank

Stata Corporation, 2001, “Svymean””, *Stata 7 Reference Manual*, Vol. 4: 52-74, College Station, Texas: Stata Press

Cục thống kê Nam Phi và Ngân hàng Thế giới, 2000 (Statistics South Africa and the World Bank, 2000), “Có phải thu nhập mà một phương pháp đánh giá đầy đủ phúc lợi xã hội đối với từng hộ gia đình : kết hợp điều tra dân số và số liệu điều tra để xây dựng bản đồ đói nghèo cho Nam Phi” (“Is census income an adequate measure of household welfare: combining census and survey data to construct a poverty map of South Africa”), Mimeo

Tobler W., 1990, Khung Phân tích sự độc lập không gian (*Frame Independent Spatial Analysis*), trang. 115-122 of M. Goodchild, ed., Sự chính xác của cơ sở số liệu không gian (*Accuracy of Spatial Data Bases*), Taylor and Francis, London.

Điều tra địa lý Mỹ USGS – GTOPO30 (United States Geological Survey USGS – GTOPO30,) <http://edcdaac.usgs.gov/gtopo30/gtopo30.html>, accessed 2003

Van de Walle, D. 2002. “Lựa chọn đầu tư nâng cấp đường nông thôn để xóa đói giảm nghèo” (“Choosing rural road investments to help reduce poverty.”) *World Development* 30 (4): 575-589.

Van de Walle, D. và Gunewardana, 2001, “Nguyên nhân gây bất bình đẳng các dân tộc thiểu số ở Việt Nam” (“Sources of ethnic inequality in Viet Nam”), *Journal of Development Economics*, Vol 65: 177-207

Ngân hàng thế giới, 2000 (World Bank, 2000), Đánh giá đói nghèo ở Panama : Ưu tiên và chiến lược xói đói giảm nghèo (*Panama Poverty Assessment: Priorities and Strategies for Poverty Reduction*), Washington DC: World Bank Country Study.

Phụ lục A: Sử dụng các biến GIS trong phân tích thống kê

Michael Epprecht, IFPRI

1. Giới thiệu chung về các biến GIS:

Phân tích các yếu tố ảnh hưởng đến đói nghèo này tập trung vào xem xét mối liên hệ giữa tỉ lệ đói nghèo cấp huyện và một loạt các biến về sinh thái nông nghiệp thông qua phân tích hồi quy. Biến phụ thuộc lấy từ Điều tra dân số và Điều tra dựa trên phân tích hồi lập bản đồ đói nghèo, còn hầu hết các biến giải thích độc lập được sử dụng trong phân tích này lấy từ dữ liệu GIS. Dưới đây là một số vấn đề chung của việc sử dụng các biến GIS trong phân tích thống kê.

Vấn đề đầu tiên trong phân tích này là thiết lập một mối liên hệ giải thích giữa người dân và môi trường sống của họ, chẳng hạn như giữa số liệu kinh tế xã hội (tỉ lệ đói nghèo ước tính của các tỉnh) và số liệu về môi trường trong GIS. Một điều quan trọng ở bước này là biến phụ thuộc là một kiểu dữ liệu về không gian và quy mô về không gian khác so với phần lớn các biến độc lập: trong khi các biến kinh tế xã hội là ở dưới dạng rời rạc đề cập đến các đơn vị hành chính hoặc các điểm, thì số liệu về môi trường là các biến liên tục. Điều này tạo ra khó khăn về mặt xây dựng phương pháp trong việc áp dụng phân tích về không gian, được gọi chung là khó khăn phân tích vùng có thể thay đổi (modifiable areal unit problem-MAUP). MAUP là vấn đề cố hữu đối với việc phân tích số liệu về không gian: vấn đề chính ở đây là các dạng khác nhau của quy mô số liệu về mặt không gian có thể dẫn đến các kết quả khác nhau, và như Heywood (1998) cho rằng: MAUP là một vấn đề từ việc áp đặt các đơn vị hành chính không gian một cách chủ quan vào các hiện tượng địa lý có tính liên tục sinh ra từ quá trình xây dựng các hình thái không gian chủ quan đó ”.

Rõ ràng, MAUP chính là một vấn đề cần được xem xét khi kết hợp số liệu kinh tế xã hội và môi trường để làm phân tích này.

Việc xây dựng các ước lượng về tỉ lệ đói nghèo của huyện có thể cho phép phân tích liên kết ở cấp huyện. Điều này có nghĩa là các biến về nông nghiệp sinh thái của huyện cần phải được tính toán. Và kết quả này một lần nữa cho thấy sự cần thiết của việc thiết lập một số vùng địa lý của không gian có tính liên tục tới các giá trị bản ghi có tính rời rạc ở các huyện. Định nghĩa giới hạn về không gian (đơn vị diện tích) của ‘huyện’ có ý nghĩa rất quan trọng đến kết quả của tập số liệu tổng hợp (aggregated data set) và cho các kết quả như sau : các biến khí hậu, nông nghiệp có thể kết hợp các số liệu kinh tế xã hội trong mỗi huyện có thể được định nghĩa là, tất cả sự quan sát trong phạm vi biên giới hành chính của một huyện, hoặc có ý

nghĩa hơn nữa là những quan sát trong phạm vi một vành đai nhất định của các huyện trong tỉnh, hoặc được coi như là một vùng là trung tâm kinh tế xã hội chính của tỉnh.

Bên cạnh đó, cần phải xem xét các tình huống sau : giả sử có hai huyện giống nhau về mặt địa lý với một vùng đồng bằng và một khu giáp ranh với núi ; một tỉnh có tới 90% dân số sống bằng nghề trồng lúa, trong khi phần lớn dân số của một tỉnh khác lại là những nông dân canh tác ở vùng cao. Trong khi các hoạt động kinh tế xã hội và sử dụng tài nguyên thiên nhiên, thực chất là mối liên hệ giữa môi trường và kinh tế xã hội, có nhiều khả năng sẽ khác biệt rất nhiều ở tỉnh A so với tỉnh B, dân số của cả hai huyện sẽ có ảnh hưởng về môi trường giống nhau.

Có thể thấy rõ là các đơn vị diện tích cho sự liên kết về dân số - môi trường có thể thay đổi và các biến tổng hợp tương ứng sẽ có giá trị ‘cấp huyện’ khác nhau, phụ thuộc vào sự lựa chọn đơn vị diện tích cụ thể.

Một mặt mỗi đơn vị diện tích như vậy có thể dễ dàng kết hợp với số liệu kinh tế xã hội của trên cùng một đơn vị diện tích, song mặt khác, thông tin về sự biến động trong vùng một đơn vị địa lý, theo định nghĩa về không gian sẽ bị mất đi. Ví dụ, một huyện có địa hình đồi núi với độ cao dao động từ 200 MaS từ đáy thung lũng đến 2700 MaS đến đỉnh núi có thể có một độ cao ý nghĩa trung bình cấp huyện là 1200 MaS, tương đương như một huyện ở cao nguyên có độ cao 1200 MaS.

Từ các kết quả này ta rút ra hai vấn đề. Thứ nhất là, việc khoanh vùng lựa chọn các đơn vị diện tích, có thể bao gồm một số loại quyền số không gian phản ánh sự phân bố dân số và các hoạt động kinh tế trong huyện (xem Epprecht & Müller, 2003) sẽ cho một phân tích tốt hơn về liên kết giữa dân số - môi trường. Thứ hai là, các biến tổng hợp khác nhau cần phải biểu thị sự biến động trong phạm vi các tỉnh một cách tốt nhất : ít nhất, bên cạnh các giá trị trung bình, các biến bổ sung phải biểu thị sự thay đổi về mặt địa lý của cùng một biến GIS (ví dụ như độ cao).

Tuy nhiên, phạm vi của dự án nghiên cứu này chỉ cho phép đánh giá vấn đề thứ hai. Đối với việc phân tích mối tương quan này, trung bình, giá trị nhỏ nhất, lớn nhất và độ lệch chuẩn của mỗi huyện được tính cho một vài dữ liệu GIS liên tục trong không gian (spatially continuous GIS data) nhằm phản ánh một vài biến động của mỗi biến trong từng huyện.

Tuy nhiên, việc ra quyết định chính sách thường dựa trên những kết quả đạt được từ phân tích thống kê mà ít nhất số liệu đã được xác định phần nào về mặt không gian (sự phân bố quỹ hỗ trợ đặc biệt cho các huyện nghèo nhất), vấn đề thứ nhất cần được lưu ý khi đưa ra kết luận từ các kết quả phân tích dưới đây và cần chú ý nhiều hơn nữa tới những vấn đề trong việc phân tích số liệu không gian sau này.

Nguồn số liệu, xử lý số liệu và việc tạo ra các biến sẽ được mô tả chi tiết hơn trong phần sau.

2. Nguồn số liệu

Các loại số liệu khác nhau và từ nhiều nguồn khác nhau được sử dụng để tạo ra các biến sử dụng trong phân tích. Các phần phân tích dưới đây sẽ cung cấp một cái nhìn tổng quát về các nguồn số liệu:

Các dữ liệu về độ cao được số hoá được sử dụng để tạo ra các biến có liên quan được dựa trên GTOPO30, một mô hình độ cao số toàn cầu (DEM) với một khoảng cách đường kẻ ô nằm ngang 30 arc giây (tương đương khoảng 1 km) theo Điều tra Địa lý của Mỹ (USGS). GTOPO30 bắt nguồn gốc từ đường quét và nguồn dữ liệu vector của thông tin về địa hình.

Dữ liệu GIS dưới dạng vector như ranh giới, đường xá, sông ngòi, vị trí địa lý của tỉnh,... được lấy từ hai nguồn chính ở Việt Nam : lớp số liệu GIS về giao thông (đường bộ và đường sắt) và hệ thống sông ngòi ; các biên giới hành chính thu được từ Trung tâm Viễn thám và Geomatics (VTGEO) tại Hà nội. Dữ liệu về đất mặt, loại đất và khí hậu lấy từ Bộ Khoa học và Công nghệ (MOSTE).

3. Xử lý số liệu và chất lượng số liệu

Trong phần này, những vấn đề về chất lượng của số liệu và độ chính xác của chúng, sự cần thiết của việc phân tích cũng như phương pháp phân tách số liệu sẽ được đề cập một cách ngắn gọn.

Trước hết ta xét đến quy mô của huyện (đơn vị địa lý qua đó số liệu được tổng hợp), sự chính xác để xác định quy mô bản đồ nguồn 1:250.000 là hợp lý. Tuy nhiên, đối với một số lớp số liệu (đất và đất mặt chẳng hạn), chỉ có số liệu dựa trên bản đồ nguồn hoặc tương tự như vậy (chủ yếu là ảnh vệ tinh), chỉ có quy mô bản đồ không lớn 1:500.000 hoặc 1:1.000.000. Số liệu độ cao của USGS có độ chính xác xấp xỉ 1:1.000.000.

Số liệu thu được từ các nguồn khác nhau được mô tả trên đây có sự khác nhau về chất lượng, do đó cần phải hiệu chỉnh. Một đánh giá chung về số liệu GIS ở Việt Nam cơ bản là tất cả các tập số liệu GIS sẵn có ở cấp quốc gia được tạo ra để áp dụng trong việc nghiên cứu và lập bản đồ. Trong khi chất lượng của số liệu nhìn chung đủ để nhìn thấy trên bản đồ, thì phần lớn sai số hình học tô pô không nhìn thấy trên bản đồ khi số liệu được sử dụng cho các mục đích phân tích về không gian và lập mô hình GIS. Sai số lớn nhất là đối với các dữ liệu đa giác và dòng kẻ số hoá. (xem chi tiết dưới đây).

3.1 Dữ liệu độ cao số hoá (Digital elevation data):

Số liệu độ cao số hoá và các số liệu khác như độ dốc, địa hình, hình dạng (trong bản đồ minh họa) được tính toán bằng cách sử dụng các nguồn số liệu đã nhắc đến ở trên với số liệu GTOPO30. Trong khi tập hợp dữ liệu, với độ chính xác tương đương sai số trung bình là căn 18 m (18 meters root mean square error), đủ để xây dựng các biến độ cao cho mỗi huyện, việc tính các giá trị các biến nhỏ nhất, lớn nhất và độ lệch chuẩn. Một quá trình phân tích

chung được xem là để bổ sung các điểm còn thiếu của tập hợp số liệu để loại bỏ những sai số của số liệu. Tương tự như vậy, giá trị độ cao nhỏ nhất và lớn nhất của tỉnh được nối từ những dữ liệu này sẽ có độ tin cậy thấp và vì thế không có trong cơ sở dữ liệu. Lớp số liệu GIS mô tả độ dốc cũng như ba định nghĩa khác về địa hình được tính toán và xây dựng trong một môi trường GIS. Độ dốc được tính toán bằng cách xác định tỉ lệ sự thay đổi lớn nhất về độ cao từ mỗi ô (grid cell) tới các ô gần kề. Địa hình được xây dựng như độ lệch chuẩn của các giá trị độ cao của tất cả các ô gần kề trong một bán kính nhất định. Bên cạnh đó, sự khác nhau về địa hình sẽ được xây dựng cho mục đích phân tích và lập bản đồ minh họa với các số liệu độ cao được số hoá.

3.2. Số liệu vector (vector line data)

Lớp số liệu đường thẳng số hoá biểu thị cho sông ngòi và mạng lưới giao thông nhìn chung đủ để tính toán các biến gần kề như ‘khoảng cách trung bình tới đường lớn’. Tuy nhiên, sai số hình học topo trong các tập hợp dữ liệu này không cho phép tính toán trực tiếp của ‘khoảng cách dọc theo mạng lưới’ như là ‘khoảng cách đường bộ tới thị trấn gần nhất’. Do đó, số liệu đường thẳng được chuyển sang số liệu dạng ô (grid cell). Biến mật độ mạng lưới vận tải được tính toán cho mỗi huyện bằng cách chồng ghép các lớp số liệu về giao thông và lớp biên giới của huyện, trước khi tính tổng chiều dài của tất cả các đường bộ trong huyện, cũng như tổng diện tích huyện. Biến mật độ được tạo ra bằng cách chia tổng chiều dài đường cho diện tích huyện..

3.3 Số liệu đa giác vector (Vector polygon data):

Lớp số liệu đa giác (polygon data layer) được sử dụng trong phân tích bao gồm chủ yếu các lớp số liệu về biên giới hành chính, loại đất và bản đồ đất mặt. Với biên giới hành chính, có hai khó khăn gặp phải. Thứ nhất, hệ thống mã hành chính phải phù hợp với hệ thống mã sử dụng trong Điều tra dân số và Nhà ở 1999, ở đó sự thay đổi các đơn vị hành chính (chia và sát nhập các đơn vị hành chính) cũng là những vấn đề không dễ. Thứ hai, một lượng lớn các sai số hình học topo (topological errors) không cho phép bất cứ sự tính toán diện tích nào trên một đơn vị hành chính. Tuy nhiên, đây là tiền đề cần thiết cho việc thiết lập nhiều biến cấp huyện. Do đó, các chức năng phân tích không gian mở rộng cần phải áp dụng phân tích tập hợp số liệu .

Mặc dù số liệu biên giới hành chính nằm trong cấu trúc dữ liệu đa chiều, tập hợp dữ liệu cấu trúc đất và loại đất mặt được chuyển sang lớp số liệu dạng ô để tính toán diện tích và phân tách dữ liệu. Trong một môi trường GIS, ‘số liệu thống kê vùng’ (trong trường hợp này là số liệu thống kê trên một vùng) được hình thành để tổng hợp các biến trên các vùng của mỗi lớp đất mặt trên mỗi vùng. Phân tích thống kê không gian tương tự được thực hiện bằng cách sử dụng lớp số liệu độ cao và các lớp dữ liệu “phái sinh” với giới hạn biên giới hành chính để tính toán giá trị các biến như giá trị trung bình, nhỏ nhất, lớn nhất và độ lệch chuẩn của độ cao, địa hình không bằng phẳng hoặc độ dốc trên mỗi vùng.

3.4 Dữ liệu Vector điểm (Vector point data):

Số liệu dạng điểm số hoá (digital point data) sử dụng trong phân tích này gồm lớp số liệu điểm vị trí các trung tâm hành chính và vị trí các trung tâm dự báo khí hậu bao gồm các bảng mô tả lượng mưa trung bình hàng tháng, nhiệt độ trung bình hàng tháng, độ ẩm trung bình hàng tháng và số giờ nắng.

Số liệu về khí hậu :

Để có thể tạo ra các biến về khí hậu cho mỗi vùng, thì các số liệu về khí hậu cần phải có đủ cho 614 huyện, chứ không phải chỉ có từ 161 trung tâm dự báo. Do các thông tin này không có đầy đủ, các nhà nghiên cứu phải áp dụng kỹ thuật nội suy không gian (spatial interpolation) để tính toán các biến khí hậu trên cả nước, từ đó mới có thể tính toán giá trị biến cho từng vùng. Để có được các biến về khí hậu chính xác, kỹ thuật lập mô hình phức tạp (sophisticated modeling techniques) sẽ được áp dụng kết hợp với các yếu tố như độ cao, loại hình đất, đất mặt,..., các yếu tố có ảnh hưởng trực tiếp đến điều kiện khí hậu của địa phương. Tuy nhiên, quá trình lập mô hình khí hậu phức tạp không thuộc phạm vi của nghiên cứu này. Do đó, phương pháp nội suy trực tiếp (straight-forward interpolation techniques) được áp dụng bằng kỹ thuật Kriging. Kriging là một phương pháp thống kê địa lý hiện đại để tạo ra bề mặt từ một tập hợp riêng rẽ các điểm với giá trị z . Không giống các phương pháp khác, giải pháp Kriging liên quan tới việc kiểm tra tác động qua lại của các hành vi không gian của hiện tượng đại diện bởi giá trị z trước khi chọn lọc phương pháp ước lượng tốt nhất để tổng hợp kết quả.

Thị xã và thành phố:

Các tập hợp số liệu này được sử dụng trong hình thái các điểm cũng như dữ liệu đa chiều. Mặc dù lớp số liệu điểm về vị trí các trung tâm hành chính trên một đơn vị hành chính (tỉnh và huyện) cần phải liên tục cập nhật sự thay đổi, việc phân loại thị xã và thành phố theo quy mô yêu cầu các đầu vào khác: Để xác định quy mô (số dân) và phạm vi diện tích của thị xã và thành phố (nhiều thị xã, và nhất là các thành phố, mở rộng diện tích ngay trong tỉnh), thông tin từ Điều tra dân số và Nhà ở 1999 được kết hợp với cơ sở dữ liệu GIS : bằng cách kết hợp thông tin về các thành thị và số dân sẵn có trong Điều tra dân số với các đơn vị hành chính tiêu biểu trong cơ sở dữ liệu GIS, ta có thể xây dựng mô hình phác hoạ về thành thị và ước tính số dân. Để tính toán khoảng cách tới thị xã hoặc thành phố, người ta tính khoảng cách tới vành đai thành thị gần nhất. Giá trị độ cao của thị xã thu được thông qua việc trùng khớp các lớp dữ liệu điểm GIS về vị trí địa lý của thị xã theo mô hình độ cao số hoá theo ô (gridded digital elevation model) và phân tách thông qua việc gán các dữ liệu theo quá trình xác định vị trí.

4. Mô tả biến:

Về mặt tổng thể, khoảng 430 biến có nguồn gốc từ lớp số liệu GIS. Bên cạnh đó, 22 biến khác về hệ thống chợ và dân số từ các kết quả của Điều tra mạng lưới chợ Việt Nam 1999 và từ Điều tra dân số và Nhà ở 1999. Dưới đây là sự tổng quan về các biến, được chia thành 6 nhóm chính :

Vị trí:

- UTM48 xác định XY theo thị trấn huyện
- UTM48 xác định XY theo trung tâm huyện

Biến kinh tế xã hội:

- Số lượng xã
- Dân số của tỉnh (tổng, nam giới, nữ giới)
- Dân số của trung tâm tỉnh lỵ
- Số lượng chợ
- Số chợ trên một xã (trung bình mỗi tỉnh)
- Thuế chợ trả cho nhà nước

Đất mặt:

- Tổng diện tích của tỉnh
- Diện tích đất trồng trọt
- Diện tích rừng tự nhiên
- Diện tích rừng trồng
- Diện tích đất trống và đất đá

Địa hình:

- Tỷ lệ tỉnh theo độ cao so với mặt nước biển từ 0-250m, 250-500m, 500-1000m, 1000-1500m, trên 1500m
- Độ cao của tỉnh (trung bình, trung vị, độ lệch tiêu chuẩn)
- Độ cao của thị trấn huyện
- Địa hình không bằng phẳng, ví dụ như độ lệch tiêu chuẩn của giá trị từng ô trong bán kính 5, 12 và 25 km (nhỏ nhất, lớn nhất, phạm vi, trung bình, độ lệch tiêu chuẩn)
- Tỷ lệ tỉnh theo độ dốc từ 0-4%, 4-8%, 8-15%, 15-30%, trên 30%
- Tỷ lệ tỉnh theo loại đất (32 loại đất)

Vận chuyển và khả năng tiếp cận :

- Độ dài của các loại đường (tổng, đường chính, đường nhỏ)
- Chiều dài của các con sông
- Khoảng cách đến đường chính (lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách đến một đường (lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách từ trung tâm huyện đến thị xã gần nhất
- Khoảng cách trung bình tới thị xã (nhỏ nhất, lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách trung bình tới thị trấn huyện (lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách từ trung tâm huyện tới trung tâm trên 10.000, 50.000, 100.000, 250.000 và 1 triệu người. (nhỏ nhất, lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách tới thị trấn trên 10.000, 50.000, 100.000, 250.000 và 1 triệu người. (nhỏ nhất, lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Khoảng cách dọc theo đường gần nhất từ trung tâm huyện tới một thành phố gần nhất trên 10.000, 50.000, 100.000, 250.000 và 1 triệu người. (nhỏ nhất, lớn nhất, trung bình, độ lệch tiêu chuẩn)
- Thời gian từ đường gần nhất tới trung tâm gần nhất trên 10.000, 50.000, 100.000, 250.000 và 1 triệu người. (nhỏ nhất, lớn nhất, trung bình, độ lệch tiêu chuẩn)

Khí hậu :

- Lượng mưa trung bình hàng tháng (nhỏ nhất, lớn nhất, chênh lệch giữa lớn nhất và nhỏ nhất, trung bình, độ lệch tiêu chuẩn)
- Nhiệt độ trung bình hàng tháng (nhỏ nhất, lớn nhất, chênh lệch giữa lớn nhất và nhỏ nhất, trung bình, độ lệch tiêu chuẩn)
- Số giờ nắng trung bình hàng tháng (nhỏ nhất, lớn nhất, chênh lệch giữa lớn nhất và nhỏ nhất, trung bình, độ lệch tiêu chuẩn)
- Độ ẩm trung bình hàng tháng (nhỏ nhất, lớn nhất, chênh lệch giữa lớn nhất và nhỏ nhất, trung bình, độ lệch tiêu chuẩn).

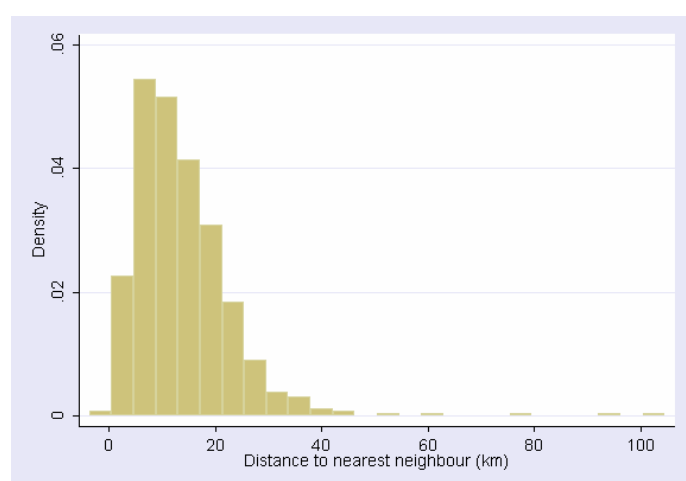
5. Ma trận quyền số không gian

Trong phân tích về số liệu không gian, một hiện tượng về tự tương quan về mặt không gian cần phải được xem xét. Sự tự tương quan về không gian là một vấn đề phụ thuộc về mặt không gian do sự giống nhau của các quan sát gần nhau. Ví dụ , sự biến động của các số liệu giống nhau thì thường gần nhau, mâu thuẫn với giả định về thống kê chung về sự độc lập của các quan sát (có thể xem chi tiết hơn ở phần dưới).

Để có thể khắc phục những tác động nhiễu trong phân tích hồi quy, mà có thể sẽ dẫn tới các kết quả sai lệch. Một cách chung nhất là xây dựng một ma trận quyền số không gian. Với một ma trận quyền số không gian, một số biện pháp được tiến hành để hạn chế tác động của các biến gần kề nhau (thực chất thì không có tiêu chuẩn nào để định nghĩa và xác định ảnh hưởng của “sự gần nhau” hay “sự tương quan lẫn nhau”). Nhìn chung, có hai loại ma trận quyền số không gian. Thứ nhất là các ma trận gần nhau một cách rời rạc với các giá trị 1 và 2 tùy thuộc vào đa giác kề nhau. Thứ hai là ma trận quyền số không gian liên tục khi các quan sát được tính theo các hàm phân tách khoảng cách (distance decay function).

Trong phạm vi của phân tích này, chúng tôi lựa chọn cách tính quyền số dựa trên khoảng cách nghịch đảo trong đó khoảng cách được tính từ trung tâm huyện. Để có thể xác định cụ thể sau đó ngưỡng của quyền số theo khoảng cách bằng 0, giới hạn khoảng cách mà trong đó các cặp địa điểm có thể coi là gần kề, có nghĩa là liên tục về mặt không gian phải được xác định. Việc lựa chọn các giới hạn về khoảng cách có ý nghĩa trực tiếp đối với mức độ của sự phân tách quyền số theo khoảng cách. Giới hạn của khoảng cách cần phải được xác định theo cách mà mỗi quan sát có ít nhất một giá trị gần kề. Đồ thị 1 cho thấy biểu đồ về khoảng cách theo đường thẳng từ trung tâm huyện tới một trung tâm gần nhất. Rõ ràng, có một số lượng nhỏ các quan sát có khoảng cách rất lớn tới một quan sát gần nhất. Không có gì ngạc nhiên khi thấy ba huyện có khoảng cách xa nhất tới các điểm gần kề là các huyện đảo. Nếu loại bỏ các huyện đảo này như Bạch Long Vĩ (Hải Phòng), Phú Quý (Bình Thuận), Côn Đảo (Bà Rịa, Vũng Tàu) cho phép phân loại giới hạn khoảng cách lớn nhất là 75 km và như vậy, một huyện sẽ có ít nhất hai điểm gần kề. Một sự phân loại giới hạn khoảng cách quá lớn sẽ cho kết quả có độ chính xác kém do ảnh hưởng của sự tự tương quan không gian giữa các vùng, trong khi đó, một giới hạn khoảng cách quá nhỏ cũng có thể tạo ra một ma trận quyền số không gian trong đó một vài quan sát sẽ không có điểm gần kề (neighbors).

Hình A1: Khoảng cách tới các trung tâm huyện gần nhất



Phụ lục B. Định nghĩa thuật ngữ

Thuật ngữ	Định nghĩa
Khoảng tin cậy	Là một khoảng trong đó chúng ta có thể tin tưởng rằng giá trị đúng của ước lượng được xác định. Ví dụ, nếu chúng ta nói khoảng tin cậy 95% cho ước lượng của tỷ lệ đói nghèo là $45\% \pm 4\%$, điều này có nghĩa là có xác suất 95% giá trị thực của tỷ lệ đói nghèo nằm trong khoảng 41 và 49 %.
Biến phụ thuộc	Những biến trong phân tích hồi quy được giải thích bằng các biến độc lập
Hệ số Gini	Một chỉ số đo lường sự bất bình đẳng, nhận giá trị từ 0 (bình đẳng tuyệt đối) và 1 (bất bình đẳng tuyệt đối). Nó được tính bằng diện tích bên trên đường cong Lorenz và dưới đường chéo, chia cho diện tích dưới đường chéo.
Mô hình hồi quy tổng thể	Một mô hình trong đó các hệ số được giả định là không đổi giữa các vùng. Ví dụ, một mô hình hồi quy tổng thể của Việt Nam giả định rằng các hệ số của vùng này là không thay đổi so với các vùng khác của Việt Nam
Phương sai không đồng đều	Là trường hợp các phương sai của các sai số trong mô hình hồi quy biến đổi trong mẫu điều tra
Biến độc lập	Những biến trong phân tích hồi quy được dùng để giải thích biến phụ thuộc. Những biến độc lập không bị ảnh hưởng bởi giá trị của những biến độc lập khác
Mô hình hồi quy vùng	Là một hình hồi quy trong đó các hệ số được phép biến đổi theo không gian. Nhiều mô hình hồi quy được chạy trên mẫu nhỏ của tập số liệu xác định theo vị trí địa phương (vùng)
P0	Một chỉ số đói nghèo bằng tỷ lệ (%) của người dân hay của hộ sống dưới mức chuẩn nghèo, được gọi là tỷ lệ nghèo đói. Tỷ lệ đói nghèo đầu người là số người sống dưới mức nghèo đói
P1	Là chỉ số nghèo đói trong đó có xem xét đến chi tiêu đầu người bình quân của người nghèo cũng như tỷ lệ người nghèo, được gọi là độ sâu của nghèo đói hay khoảng cách đói nghèo
P2	Là một chỉ số nghèo đói, trong đó có xem xét mức độ bất bình đẳng giữa những người nghèo cũng như chi tiêu bình quân trung bình của người nghèo và tỷ lệ người nghèo, được gọi là độ trầm

	trọng của nghèo đói hay khoảng cách đói nghèo bình phương.
Đói nghèo	Là tình trạng thiếu hụt trong đó người dân không đủ nguồn lực để định hướng một cuộc sống lành mạnh và chủ động. Trong nghiên cứu này, đói nghèo được xác định qua giá chỉ tiêu bình quân đầu người, bao gồm cả giá trị lương thực ở mức đủ ăn và tiền thuê nhà ở.
Lập bản đồ đói nghèo	Một nghiên cứu ước lượng tỷ lệ đói nghèo cho khu vực nhỏ và trình bày các kết quả lên bản đồ. Thường, điều này được thực hiện bằng cách kết hợp điều tra hộ gia đình và Tổng điều tra dân số và nhà ở.
Phân tích hồi quy	Một phân tích thống kê xác định các phương trình mô tả tốt nhất tương quan giữa một hay nhiều biến phụ thuộc với một tập hợp biến độc lập.
Ước lượng diện tích nhỏ	Một phương pháp kết hợp số liệu điều tra và Tổng điều tra để tạo ra những ước lượng cho biến trong những khu vực nhỏ. Những biến quan tâm (ví dụ như đói nghèo) là những biến có trong số liệu điều tra hộ nhưng không có trong Tổng Điều tra.
Tự tương quan về không gian	Là tình trạng trong đó các giá trị của một biến ở một vùng có tương quan với giá trị của biến ở những vùng lân cận.
Sự phụ thuộc sai số về không gian	Là tình trạng trong một phân tích hồi quy khi các phần dư (sai số giữa giá trị ước lượng và giá trị thực của biến phụ thuộc) có tự tương quan về không gian.
Sự phụ thuộc trễ về không gian	Là tình trạng trong phân tích hồi quy khi biến phụ thuộc ở một vùng bị ảnh hưởng bởi giá trị của các biến ở vùng lân cận.
Ma trận quyền số không gian	Một ma trận hay tập hợp mô tả “sự gần kề” của từng cặp quan sát trong tập số liệu. Ví dụ, nhân tố (i,j) trong ma trận mô tả sự “gần kề” của quan sát i và j . “Sự gần kề” thường được xác định dựa trên khoảng cách giữa hai quan sát hay dựa trên việc xem xét hộ có liên nhau hay không
Sai số chuẩn	Một chỉ số về mức độ chính xác của các ước lượng thống kê. Dưới những điều kiện nhất định, khoảng tin cậy 95% bằng hai lần sai số chuẩn
Theil L	Một chỉ số về bất bình đẳng, được gọi là $GE(0)$, là một chỉ số đói nghèo trong tập hợp các chỉ số entropy tổng hợp
Theil T	Một chỉ số về bất bình đẳng, được gọi là $GE(0)$, là một chỉ số đói nghèo trong tập hợp các chỉ số entropy tổng hợp

