

PHỤ LỤC

PHỤ LỤC 1: TÍNH TOÁN TSGTKS VÀ KÍCH THƯỚC MẪU

Tỷ số giới tính khi sinh được tính bằng số trẻ em trai được sinh ra trên 100 trẻ em gái. TSGTKS ở mức chuẩn sinh học là 105, nhưng số liệu thống kê trên phạm vi toàn cầu cho thấy thực tế tỷ số này dao động trong khoảng 104-106, với mức độ biến thiên nhỏ ở mỗi quốc gia và theo thời gian. Khi tính toán TSGTKS trên số mẫu lớn, không bị sai số từ hệ thống đăng ký khai sinh tin cậy, mức độ sai số duy trì ở mức 1 điểm phần trăm. Điều đó có nghĩa là sai số không thể là nguyên nhân của hiện tượng gia tăng TSGTKS quan sát thấy ở nhiều khu vực châu Á, bao gồm cả Việt Nam. Tuy nhiên, mức độ nhạy cảm của tính toán TSGTKS với cỡ mẫu nhỏ là một vấn đề đã được biết đến trong phân tích TSGT. Bảng 6 cho thấy độ nhạy cảm của tính toán thông qua khoảng tin cậy 95% của các ước lượng cho các mẫu từ 1.000 đến 100.000 quan sát. Các tính toán này dựa trên mức TSGTKS ước lượng hiện hành của Việt Nam (mức 110,6). Có thể thấy mức 110,6 không khác biệt có ý nghĩa so với mức bình thường 105 với mẫu quan sát 5.000 ca. Với mẫu 20.000 ca sinh thì khoảng tin cậy tương đương với mức +/- 3 phần trăm. (107,6-113,7)

Bảng 6: Các ước lượng TSGT theo kích thước mẫu với khoảng tin cậy 95%

Kích thước mẫu	1 000	2 000	5 000	10 000	20 000	50 000	100 000	200 000	500 000
TSGTKS	110,6	110,6	110,6	110,6	110,6	110,6	110,6	110,6	110,6
Ngưỡng dưới của ước lượng	97,7	101,3	104,6	106,3	107,6	108,7	109,2	109,6	110,0
Ngưỡng trên của ước lượng	125,3	120,8	116,9	115,0	113,7	112,6	112,0	111,6	111,2

Như vậy, kích thước mẫu rất quan trọng đảm bảo kết quả có ý nghĩa thống kê của TSGTKS. Bảng 7 trình bày đặc điểm của mẫu và các tập hợp con được sử dụng cho phân tích. Cột thứ nhất cho biết số quan sát, cột thứ hai cho biết ước lượng đã được gia quyền (các số thập phân được làm tròn). Dòng cuối chỉ số sinh gần nhất và một số hạn chế trong mẫu. Ước lượng của các tình hình bày trong Phụ lục 4.

Bảng 7: Mẫu sử dụng cho các ước lượng khác nhau, số liệu mẫu TĐTDS 2009

	Mẫu quan sát	Mẫu đã gia quyền
Số hộ gia đình	14.177.590	14.177.590
Số cá nhân	3.692.042	3.692.042
Tỉnh: Dân số trung bình	225.041	225.041
Tỉnh có dân số lớn nhất	663.169 (Hà Nội)	1.177.203 (TP. Hồ Chí Minh)
Tỉnh có dân số ít nhất	117.367 (Lai Châu)	61.168 (Lai Châu)
Khu vực đô thị	3.621.262	4.193.352
Dân tộc ít người	3.281.357	2.014.438
Dân số dưới 5 tuổi	1.260.187	1.202.904
Phụ nữ tuổi 15-49	4.021.751	4.053.178
Số phụ nữ sinh con trong 12 tháng trước	260.768	247.632
Số sinh trong 12 tháng trước	262.232	247.603

PHỤ LỤC 2: SỬ DỤNG SỐ SINH BÁO CÁO ĐỂ TÍNH TOÁN TSGTKS

TSGTKS được tính toán dựa trên số sinh báo cáo trong 12 tháng trước khi diễn ra cuộc điều tra. Tuy nhiên, một số ca sinh gần đây không được đưa vào phân tích vì các lý do khác nhau²¹. Ngoài ra, cũng cần nhấn mạnh rằng, không thể sử dụng chỉ báo này để đánh giá xu hướng gần đây về mức độ gia tăng TSGTKS bằng cách tính toán các mức TSGTKS cho các giai đoạn trước. Lý do là vì TSGTKS của “lần sinh cuối” có xu hướng bị chệch do tâm lý ưa thích con trai như trình bày trong Bảng 8. Sai số tăng TSGTKS ở các lần sinh xảy ra đã lâu trong quá khứ thường là do sự gia tăng tỷ lệ những “lần sinh cuối” hay “lần sinh gần nhất” được báo cáo mà thực tế là “lần sinh cuối cùng”. Tuy nhiên do nhiều cặp vợ chồng quyết định ngừng sinh sau khi đã có con trai, dẫn đến sự gia tăng đáng kể số con trai trong các lần sinh cuối cùng. Ví dụ, TSGTKS của “lần sinh cuối” năm 2000 là 128,8 một mức cao bất thường so với mức TSGTKS thực tế. Việc bỏ qua các lần sinh con gái cũng là một yếu tố làm sai số. Tuy nhiên không có sai số nào khi tính toán TSGTKS trong 12 tháng trước.

Bảng 8: TSGT của “lần sinh cuối” theo năm sinh, số liệu mẫu TĐTDS 2009

Năm	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
TSGTKS	128,8	123,8	122,2	119,6	117,0	116,3	114,4	112,9	110,4	112,4

²¹ Trong trường hợp các ca sinh của các phụ nữ đã tử vong hoặc di cư trước cuộc điều tra bị ghi thiếu trong mẫu, nhưng điều này không tạo nên sự khác biệt trong các ước lượng.

Lần sinh cuối báo cáo trong TĐTDS có thể có liên hệ với nhiều đặc điểm của bà mẹ cũng như hộ gia đình. Tuy nhiên rất khó kết nối TSGTKS với các đặc điểm của người cha: không phải chỉ do một số bà mẹ không chung sống với chồng (do di cư, li hôn hoặc góa), mà còn do cách hỏi của TĐTDS không cho phép xác định ai là chồng của mỗi phụ nữ trong gia đình²².

Khi số sinh của lần sinh cuối quá nhỏ, TSGTTE được sử dụng và tính toán cho nhóm trẻ em dưới 5 tuổi. Nhóm dân số này có cỡ mẫu gấp 5 lần số ca sinh trong 1 năm trước (vì vậy cung cấp khoảng tin cậy cao gấp khoảng 2 lần so với lần sinh cuối). Tuy nhiên, chỉ số này bị ảnh hưởng bởi sự khác biệt về mức tử vong trẻ sơ sinh và trẻ em theo giới tính (mức tử vong trẻ em trai về sinh học cao hơn một chút so với trẻ em gái). Ngược lại với đo lường dựa trên số sinh các bà mẹ khai báo, TSGTTE không thể kết nối được với đặc tính xã hội và nhân khẩu học của bà mẹ.

PHỤ LỤC 3: TẠI SAO TSGTKS KHÁC VỚI TSGTTE DƯỚI MỘT TUỔI?

So sánh các TSGT rút ra từ số liệu TĐTDS 2009 cho thấy sự khác biệt nhỏ giữa TSGT trong vòng 12 tháng trước thời điểm điều tra (110,6) (dựa trên số sinh do các bà mẹ khai báo) và TSGT trẻ em từ 0-11 tháng tuổi (112,6). Với số lượng quan sát của hai nhóm (247.603 ca sinh và 251.744 trẻ sơ sinh) thì sự khác biệt TSGT này là đáng kể và cần làm rõ. Nhóm trẻ em dưới 1 tuổi được sinh ra trong 12 tháng trước thời điểm điều tra nhưng bị tác động của mức chết, về sinh học vốn cao hơn ở trẻ em trai. Vì vậy, thông thường TSGTTE dưới 1 tuổi hơn thấp hơn TSGTKS, chứ không phải cao hơn như trong ước lượng của báo cáo này. Tương tự như vậy, TSGTTE dưới 5 tuổi (115,5), là nhóm dân số sinh ra trong giai đoạn 2004-2009, cao hơn đáng kể so với TSGTKS của năm 2009²³.

Khả năng báo cáo thiếu số sinh trong 12 tháng trước hoặc số trẻ em dưới một tuổi cũng được xem xét bằng cách so sánh số trẻ em và số sinh trong mỗi hộ gia đình. Kết quả là số trẻ em và số sinh trong từng hộ gia đình khớp nhau hoàn toàn, không có sự khác biệt nào đáng kể về giới. Như vậy sự khác biệt giữa TSGTKS và TSGTTE dưới 1 tuổi là do hệ thống quyền số theo giới áp dụng cho các cá nhân. Quyền số mẫu gốc được thiết kế cho các đơn vị lấy mẫu và phân tầng điều tra. Tuy nhiên, theo các báo cáo của TĐTDS, hệ thống quyền số sau đó được điều chỉnh cho khớp với tổng dân số theo giới tính năm 2009. Kết quả là quyền số không chỉ biến thiên theo hộ gia đình mà còn theo giới tính. Thực tế là quyền số cho con trai cao hơn quyền số cho con gái *ngay trong cùng một đơn vị lấy mẫu và trong cùng một hộ gia đình*.

Tác dụng của phương pháp gia quyền điều chỉnh đối với giới tính phân bổ theo tuổi được minh họa trong Hình 20 trình bày quyền số trung bình theo giới tính của trẻ em dưới 15 tuổi trong toàn bộ mẫu điều tra²⁴. Có thể thấy quyền số cho nam giới cao hơn nữ giới ở tất cả các nhóm tuổi. TSGT của các quyền số này trình bày bên phải của Hình 20 nằm trong khoảng 102-104, luôn lớn hơn 100. Nói cách khác, trong cùng một quận/huyện điều tra hoặc trong

²² Trường hợp duy nhất có thể xác định người cha là các hộ gia đình hạt nhân, trong đó cha mẹ cũng là chủ hộ (có biến số xác định quan hệ với chủ hộ). Trong mẫu nghiên cứu, số này chiếm không đến 57% số phụ nữ có sinh con trong năm trước điều tra.

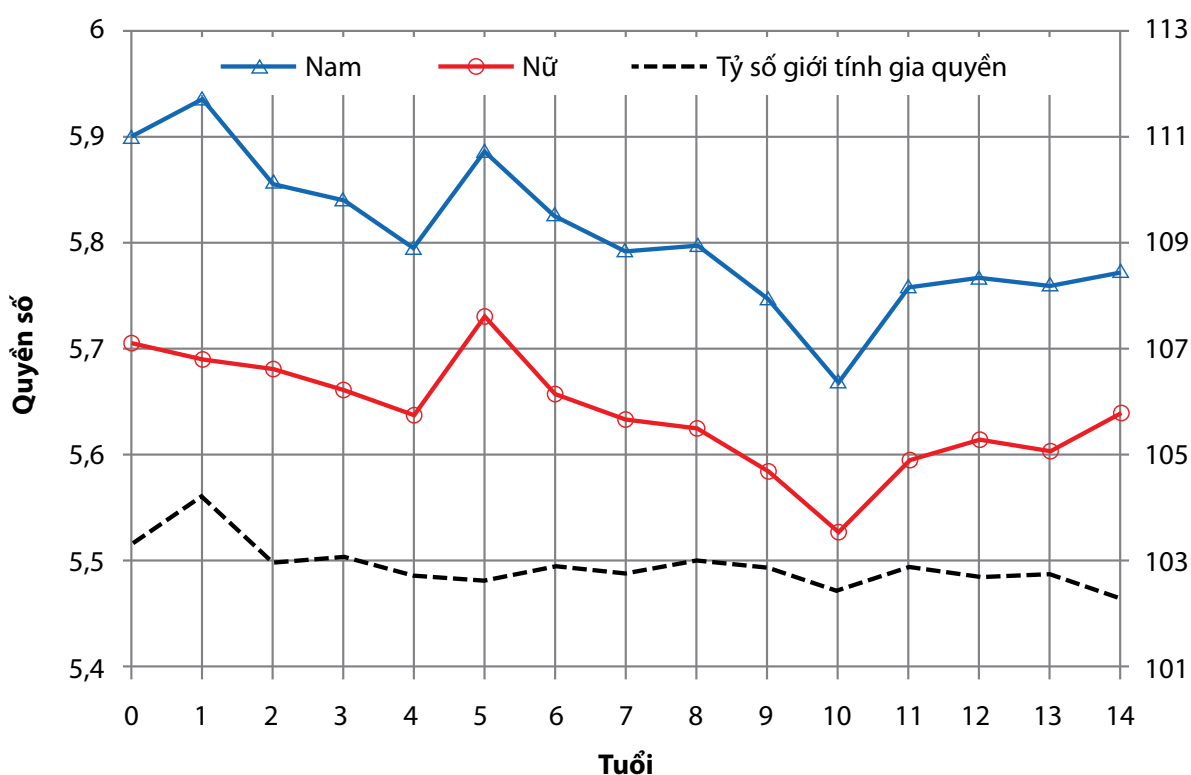
²³ Sự khác biệt do di cư theo giới tính ở lứa tuổi này không đáng kể.

²⁴ Giá trị trung bình này thu được từ toàn bộ mẫu. Ngoài sự khác biệt theo giới tính do quyền số, những số liệu cũng bị tác động của phân bổ của các đơn vị lấy mẫu theo tuổi.

cùng một hộ gia đình, con trai nhận được gia quyền với quyền số thống kê cao hơn con gái. Sự khác biệt này tác động trực tiếp đến TSGT của toàn bộ dân số. Hơn nữa, quyền số thay đổi theo tuổi, tăng lên cho nhóm 5 và giảm cho nhóm 10 tuổi. Cũng có những thay đổi riêng biệt cho nhóm 1 tuổi: trong khi quyền số cho trẻ em gái giảm dần từ 0 đến 4 tuổi, quyền số cho trẻ em trai lại tăng bất ngờ ở nhóm 1 tuổi. Kết quả là, TSGT của nhóm dân số 1 tuổi cao bất thường (112,6) và cao hơn TSGT của các nhóm trẻ em khác. Điều này tác động đến việc tính toán TSGT dựa trên phân bố tuổi và giới tính và giải thích tại sao TSGT của nhóm trẻ em 1 tuổi lại cao bất thường (112,6).

Để khắc phục những khó khăn liên quan đến quyền số theo giới tính, trong một số trường hợp có thể sử dụng quyền số hộ gia đình – không đặc trưng theo giới tính – thay vì dùng quyền số cá nhân để tính toán TSGT. Quyền số hộ gia đình được dựa trên quyền số của huyện điều tra trong thiết kế dàn mẫu và sau đó điều chỉnh loại bỏ tình trạng mất cân bằng giữa mẫu của các vùng. Sử dụng quyền số hộ gia đình, TSGT của trẻ em dưới 1 tuổi là 109,7. Do mức chết trẻ sơ sinh ở trẻ em trai về sinh học cao hơn trẻ em gái, TSGTTE cho trẻ sơ sinh được điều chỉnh mức chết trẻ sơ sinh, kết quả cho TSGTKS là 110,7 cho trẻ em dưới 1 tuổi theo số liệu TĐTDS²⁵. Giá trị này sát với mức TSGTKS 110,6, dựa trên số ca sinh báo cáo và thể hiện tính nhất quán trong các số liệu đưa ra. Mặc dù vậy, chỉ có số liệu đầy đủ của TĐTDS tính cho toàn bộ dân số mới cung cấp chính xác sự phân bố theo tuổi và giới tính trên phạm vi cả nước.

Hình 20: Quyền số thống kê theo tuổi, giới tính, và TSGT tương ứng



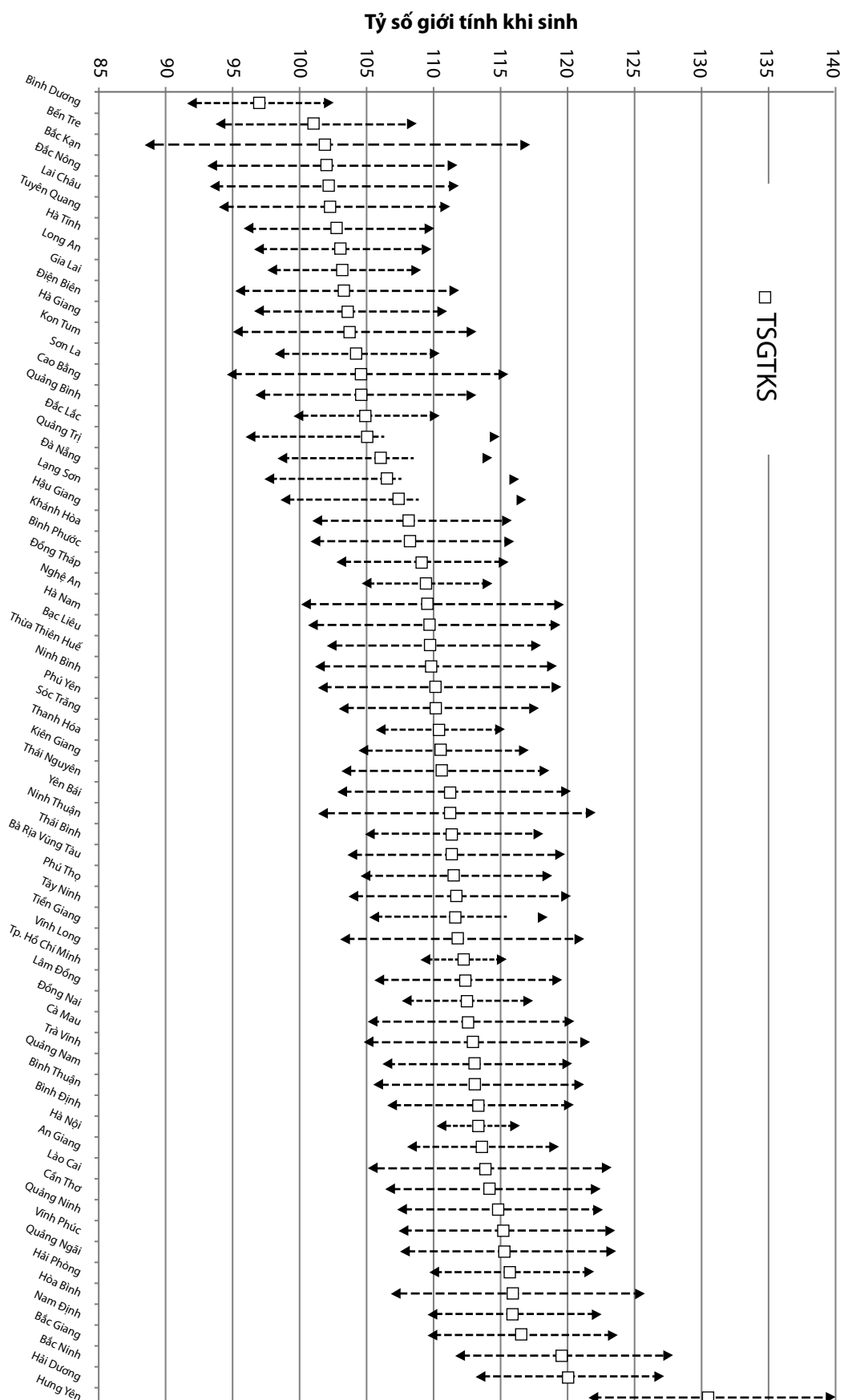
²⁵ Sử dụng Bảng sống của TĐTDS theo giới tính để dự báo ngược từ sơ sinh sang số sinh (TĐTDS 2010).

PHỤ LỤC 4: TSGTKS THEO TỈNH

Hình 21 trình bày TSGTKS theo tỉnh dựa trên số sinh trong 12 tháng trước khi điều tra. Như đã đề cập ở trên, số ca sinh thường khá nhỏ ở mỗi tỉnh và mẫu chỉ bao gồm 15% toàn bộ dân số. Số ca sinh trung bình cho mỗi tỉnh trong năm trước điều tra là 3.900. Với cỡ mẫu này, sai số ngẫu nhiên khi tính toán TSGTKS là khó tránh khỏi. Trong điều tra mẫu 15% của TĐTDS 2009, khoảng 54 tỉnh có số mẫu dưới 5.000 ca sinh.

Hình 21 cũng trình bày TSGTKS của các tỉnh với khoảng tin cậy 95%. Như dự đoán, khoảng tin cậy của các ước lượng thường lớn hơn 15% giá trị của ước lượng là do số ca sinh nhỏ. Số liệu cho thấy hơn một nửa số tỉnh có TSGTKS không khác biệt so với mức sinh học 105. Khi so sánh với mức trung bình cả nước (110,6), chỉ có 3 tỉnh có tỷ số thấp hơn, và 4 tỉnh có tỷ số cao hơn mức này. Những ước lượng này cho thấy khả năng nhiều là do sai chệch ngẫu nhiên và sai số trong đo lường.

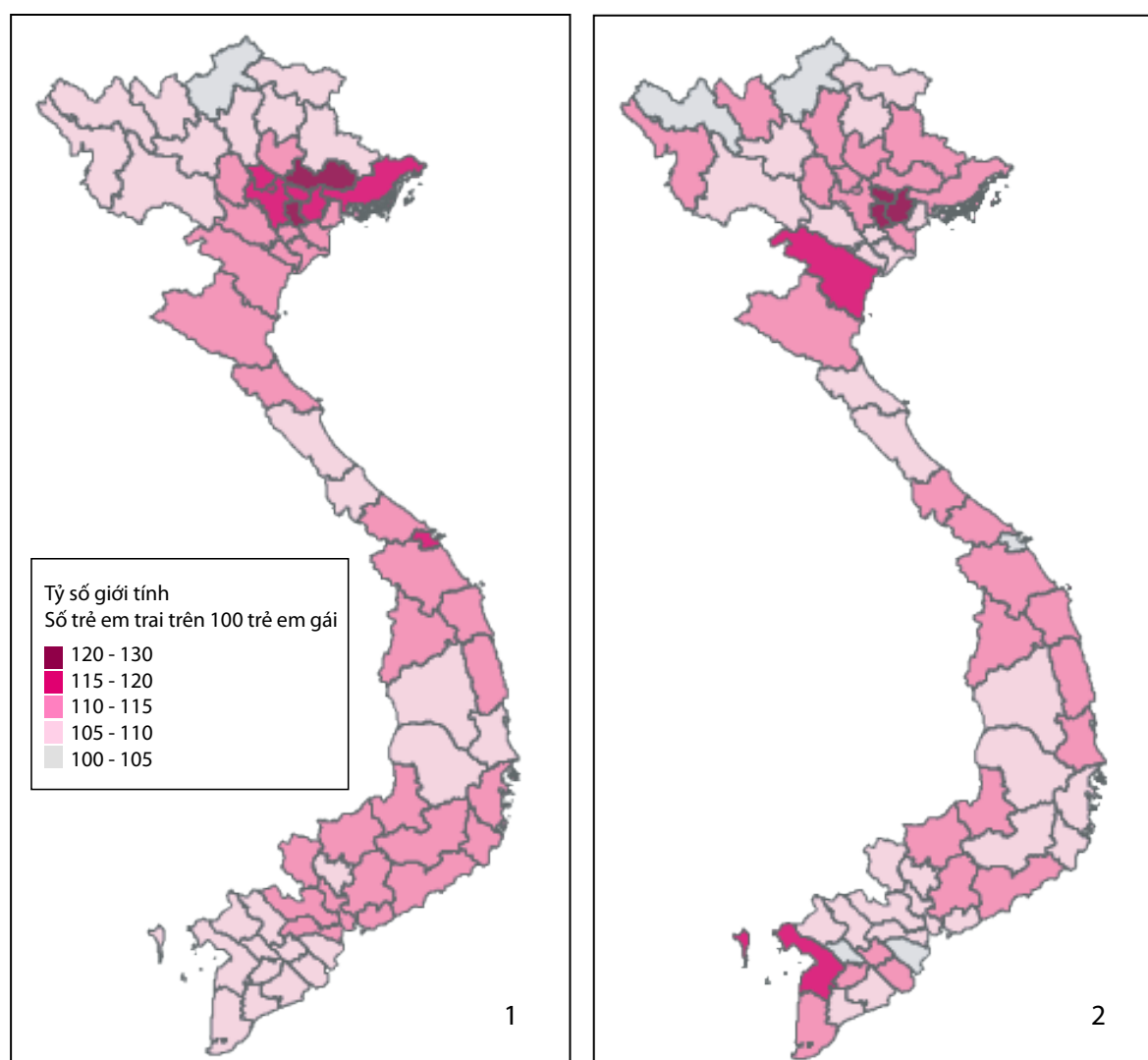
Hình 21: TSGTKS của các tỉnh với khoảng tin cậy 95%, mẫu TĐTDS 2009



PHỤ LỤC 5: PHÂN TÍCH KHÔNG GIAN VỀ SỰ BIẾN THIÊN CỦA TSGT

Phần này phân tích sâu hơn xu hướng phân bố không gian mức độ biến thiên của TSGT theo các tỉnh. Hình 22 là bản đồ TSGT cấp tỉnh dựa trên các nguồn số liệu khác nhau. Bản đồ thứ nhất trình bày TSGTTE dưới 5 tuổi theo mẫu TĐTDS 2009, bản đồ thứ hai là TSGTKS dựa trên số ca sinh rút ra từ thống kê các ca sinh từ các cơ sở y tế (trung bình cho giai đoạn 2007-2009). Hai bản đồ này không hoàn toàn khớp nhau ($r = 0,58$), và có một số điểm khác nhau rõ rệt (như Huế và Kiên Giang). Tuy nhiên, cả hai đều phản ánh một xu hướng chung, với một cụm các tỉnh ở vùng Đồng bằng sông Hồng có mức TSGT cao hơn hẳn các vùng miền Trung và miền Nam.

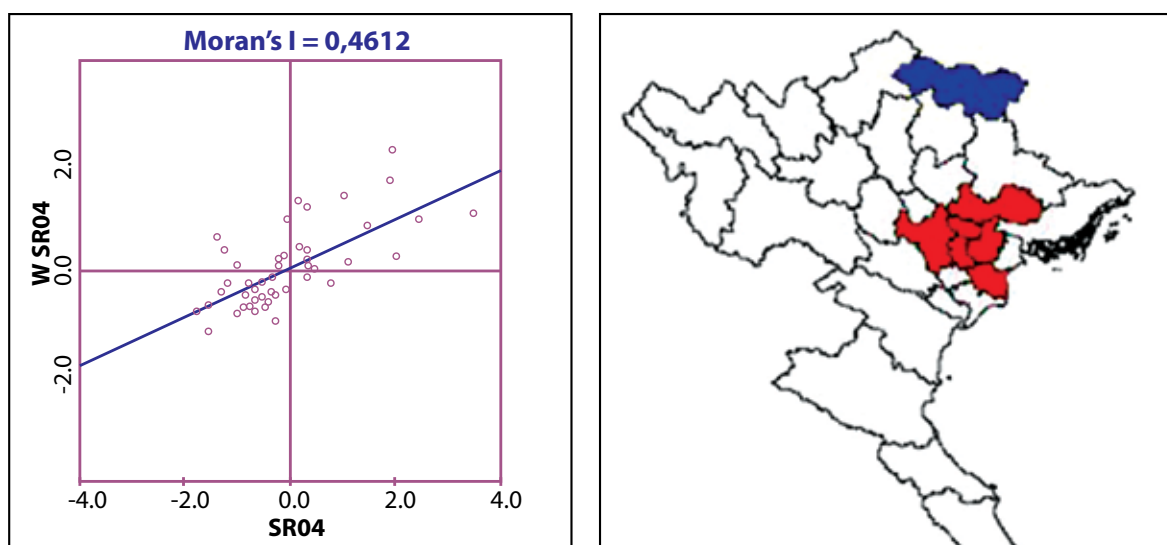
Hình 22: TSGTKS theo các tỉnh: 1. TSGT trẻ em 0-4 tuổi năm 2009 và 2. TSGTKS trung bình 2007-2009



Để đánh giá sự phân bố không gian của TSGT ở Việt Nam, hai bản đồ trên (bản đồ TSGTTE 0-4 tuổi 2009 và bản đồ TSGTKS 2007-2009) được sử dụng cùng với bản đồ TSGTKS trong 12 tháng trước thời điểm điều tra. Mỗi bản đồ được tiến hành phân tích thống kê địa lý nhằm đánh giá mức độ tương quan không gian và các chỉ báo về TSGT theo các cụm. Việc lập bản đồ và tính toán tiến hành trên phần mềm Geoda (Anselin 2003).

Bản đồ TSGTTE cho thấy mức độ tương quan mạnh mẽ nhất, được đo bằng chỉ số Moran I, kết quả trình bày trong Hình 23. Hình bên trái (phân bố Moran) cho thấy TSGTTE theo các tỉnh năm 2009 so với TSGTTE trung bình của các tỉnh lân cận. Mức độ phụ thuộc rất mạnh và tương ứng với mức tương quan không gian (Moran's I) là 0,46. Hình bên phải là bản đồ miền Bắc Việt Nam với chỉ số Moran cho sáu tỉnh (màu đỏ) có TSGT cao: Hà Nội, Bắc Ninh, Hưng Yên, Hải Dương, Bắc Giang. TSGT của các tỉnh này và các tỉnh lân cận chúng cao hơn hẳn mức trung bình. Những tỉnh này cũng là những vùng nóng (hot spot) về tình trạng mất cân bằng giới tính trên cả nước, và được bao bọc bởi các tỉnh lân cận (như là Vĩnh Phúc, Hoà Bình, Hải Phòng) tạo thành các khu vực địa giới. Điểm nổi bật là nhóm sáu tỉnh ($p < 0,01$) được xác định trong mô phỏng chỉ số vùng của tương quan không gian (local indicator of spatial association - LISA). Bản đồ cũng chỉ ra rằng tỉnh Cao Bằng (màu xanh) và các tỉnh xung quanh, sát biên giới Trung Quốc có TSGT thấp hơn hẳn.

Hình 23: Phân bố Moran về TSGTTE và cụm không gian, mẫu TĐTDS 2009



PHỤ LỤC 6: PHƯƠNG PHÁP THIẾT LẬP LẠI CƠ CẤU GIA ĐÌNH

Một hạn chế trong phân tích của báo cáo này là không có thông tin về sự ảnh hưởng của giới tính của lần sinh trước đến TSGTKS. Hạn chế này là do TĐTDS không thu thập thông tin về thứ tự sinh và giới tính, vì vậy chỉ có tổng số lần sinh mà không có thông tin về giới tính. Những ước lượng hồi cứu cũng không thể thực hiện do sai số về phía con trai đối với những lần sinh cuối cùng. Kết quả là không thể xác định chắc chắn lần sinh trước là trai hay gái. Đây là hạn chế lớn trong phân tích vì một giả thuyết quan trọng của nghiên cứu này về khung lý thuyết của thiên lệch giới dẫn đến chọn lọc giới tính là các cặp vợ chồng không có con trai có xu hướng phân biệt đối xử mạnh mẽ hơn với các thai gái so với những cặp vợ chồng đã có con trai. Để bổ sung, phân tích sử dụng phương pháp “con ruột” (own-child method). Đây là phương pháp thiết lập lại cấu trúc gia đình, được phát triển nhằm sử dụng cơ cấu chi tiết hộ gia đình để ước lượng những lần sinh trong quá khứ (Cho và các cộng sự, 1986). Phương pháp này nhằm xác định các bà mẹ và trẻ em trong cùng một hộ gia đình để ước lượng những lần sinh trong quá khứ, bằng cách sử dụng phân loại trẻ em theo tuổi từ số liệu điều tra (sau khi đã điều chỉnh theo mức chết). Phương pháp này được sử dụng để xác định cha mẹ và trẻ em trong mỗi hộ gia đình bằng cách sử dụng các thông tin về giới tính, tuổi, tình trạng hôn nhân, và gia đình (quan hệ với chủ hộ). Khi thông tin của một người mẹ hoặc cha có thể liên kết được với các con của họ, có thể xác định chính xác gia đình hạt nhân và sau đó xếp thứ tự trẻ em theo tuổi để xác định thứ tự sinh. Quy trình áp dụng ở đây dựa trên các thông tin gia đình sau:

1. Trẻ em là “con của chủ hộ”
2. Trẻ em là “cháu của chủ hộ” và chủ hộ chỉ có 1 người con đã lập gia đình sống cùng nhà (thường là cha).
3. Trong các trường hợp khác, trẻ được xác định có thể không phải là anh em ruột, do vậy không đưa vào phân tích

Trong hai trường hợp đầu, tương ứng với phần lớn số trẻ (95%), quy trình tái phân loại này giúp xác định các nhóm anh chị em ruột. Trong mỗi nhóm này, có thể sử dụng thông tin về ngày sinh để xác định thứ tự sinh²⁶. Ngoài ra có thể chỉ ra từ “nhóm quần thể anh chị em ruột” này là những đứa trẻ này đã có anh trai hay không khi chúng được sinh ra. Những thông tin khác có thể kiểm tra là năm sinh, tình trạng kinh tế xã hội của hộ gia đình, vùng định cư. Số liệu sử dụng tập trung vào số sinh trong 10 năm trước điều tra (quần thể dân cư sinh ra sau TĐTDS năm 1999) để tránh những sai số liên quan đến tình trạng không có anh em ruột. Do sai số có thể xảy ra do hệ thống quyền số đã mô tả ở trên, *quyền số mẫu hộ gia đình* được sử dụng thay cho quyền số mẫu cá nhân.

²⁶ Sự khác biệt chủ yếu giữa thứ tự con trong gia đình và thứ tự sinh là do mức chết và di cư. Thực tế là thứ tự con trong gia đình thường thấp hơn thứ tự sinh. Điều này sẽ tăng dần theo tuổi vì tỷ lệ lớn hơn anh chị em ruột có thể không có mặt tại hộ gia đình trong thời gian điều tra do đã kết hôn, di cư hoặc chết.

PHỤ LỤC 7: XÂY DỰNG CHỈ SỐ TỔNG HỢP VỀ TÌNH TRẠNG KINH TẾ-XÃ HỘI

Phân tích về sự khác biệt theo tình trạng kinh tế-xã hội trong trường hợp lý tưởng đòi hỏi phải có các ước lượng về thu nhập của hộ gia đình hoặc phân loại chi tiết nghề nghiệp xã hội của lực lượng lao động. Nhưng TĐTDS không thu thập thông tin của cả hai góc độ này. Thay vào đó, các thông tin hộ gia đình được sử dụng để tạo ra một thước đo tình trạng kinh tế-xã hội của hộ gia đình.

Trước tiên, một nhóm các câu hỏi về hộ gia đình liên quan đến tình trạng kinh tế xã hội được xác định và sử dụng kỹ thuật phân tích nhân tố để tính toán một hệ số tổng hợp ban đầu về tình trạng kinh tế-xã hội. Trong trường hợp này kỹ thuật phân tích tương ứng đa biến (*multiple correspondence analysis - MCA*) được áp dụng cho các biến hộ gia đình và nhà ở. Khoảng 10% toàn bộ mẫu điều tra được sử dụng trong phân tích này. Các quan sát thiếu câu trả lời (0,9%) được loại bỏ vì vậy bộ số liệu mới có cỡ mẫu nhỏ hơn so với bộ số liệu gốc. Sau lần phân tích đầu tiên, những biến số có ít quan hệ với yếu tố đầu tiên trong phân tích MCA cũng được loại bỏ.

Hệ số (hoặc nhân tố thành phần) đầu tiên rút ra từ phân tích MCA này đại diện cho 73,1% phương sai tổng thể. Đây là một giá trị rất cao khi xem xét trên góc độ có rất nhiều các biến số khác trong phân tích (xem bảng dưới). Tuy nhiên điều này tương ứng với việc chỉ có một hệ thống phân loại tình trạng kinh tế-xã hội duy nhất để xếp hạng các hộ gia đình. Hệ số đầu tiên, dao động quanh giá trị trung bình 0, được tổng hợp từ các biến sau: sở hữu 7 loại tài sản khác nhau của hộ gia đình (TV, điện thoại, máy tính, máy giặt, tủ lạnh, điều hoà, xe máy), 4 loại tiện nghi trong nhà (chiếu sáng, nhiên liệu sử dụng cho nấu ăn, nước uống, và nhà vệ sinh), và vật liệu xây dựng nhà (bao gồm tường và mái nhà). Tất cả các biến số gốc này đều có liên quan chặt chẽ - thuận chiều hay ngược chiều - với hệ số tổng hợp hộ gia đình. Cuối cùng, hệ số mức sống này được sử dụng để phân loại số sinh theo 5 nhóm kinh tế-xã hội, bắt đầu từ các hộ gia đình thuộc 20% nghèo nhất đến 20% giàu nhất²⁷.

Bảng 9 minh họa tính nhất quán của hệ số này, theo đó, phân bố phụ nữ theo trình độ giáo dục được tính theo nhóm kinh tế-xã hội. Số liệu trong Bảng 9 cho thấy mối liên quan chặt chẽ giữa trình độ giáo dục của phụ nữ với trình trạng kinh tế-xã hội hộ gia đình. Thậm chí ngay cả khi hai biến số này hoàn toàn không trùng khớp (ví dụ: một số hộ gia đình thuộc nhóm giàu nhất có một tỷ lệ khá lớn phụ nữ không được đi học) thì trình độ giáo dục của phụ nữ và mức sống hộ gia đình vẫn thể hiện mối tương quan rất chặt chẽ. Các số liệu khác cũng cho thấy mối tương quan chặt chẽ giữa tình trạng kinh tế-xã hội và các biến số khác như khu vực nông thôn/thành thị hay dân tộc.

Bảng 9: Trình độ giáo dục của bà mẹ theo nhóm kinh tế-xã hội, mẫu TĐTDS 2009

Nhóm kinh tế xã hội	Trình độ giáo dục tiểu học	Trình độ giáo dục trung học phổ thông	Trình độ giáo dục cao hơn	Tổng số
Nghèo nhất	50,4%	49,2%	0,4%	100,0%
Nghèo	32,9%	66,0%	1,1%	100,0%
Trung bình	17,8%	79,5%	2,7%	100,0%
Giàu	12,9%	79,9%	7,3%	100,0%
Giàu nhất	4,8%	63,6%	31,6%	100,0%
Tổng số	18,5%	70,0%	11,4%	100,0%

²⁷ Những phân nhóm nhỏ hơn dẫn đến số sinh quá nhỏ cho phép có phân tích thống kê vững.

PHỤ LỤC 8: DỰ BÁO DÂN SỐ

Các dự báo dân số đến năm 2049 được trình bày trong báo cáo này là các mô phỏng theo các tham số nhân khẩu học (mức sinh, mức chết) từ các dự báo gần đây của Vụ Dân số của Liên hợp quốc (2009), với cấu trúc tuổi và giới từ TĐTDS 2009 được dùng làm số liệu gốc. Các dự báo này tương tự như các dự báo khác, với khác biệt duy nhất là các yếu tố trong từng mức độ khác nhau của TSGTKS được đưa vào để mô phỏng tác động của chúng lên phân bố giới tính tương lai.

Bảng 10: Các tham số dự báo, 2009-2049

Năm	2009	2015	2020	2030	2049
Tổng tỷ suất sinh	2,1	1,9	1,9	1,85	1,85
Mức chết nam giới (tuổi thọ bình quân)	70,3	71,6	72,6	74,5	79,0
Mức chết nữ giới (tuổi thọ bình quân)	76,2	77,2	78,0	79,5	83,0
TSGTKS: Tình huống 1	111	115	115	115	115
TSGTKS: Tình huống 2	111	113	115	105	105
TSGTKS: Tình huống 3	105	105	105	105	105
Lưu ý: Giá trị trung bình được sử dụng					

Bắt đầu với TSGTKS ở mức 111 cho năm 2009, báo cáo đưa ra hai tình huống dự báo về TSGTKS khác nhau. Tình huống thứ nhất (không có can thiệp), TSGTKS đạt 115 vào năm 2015 và ổn định ở mức này trong những năm tiếp theo. Tình huống thứ hai lạc quan hơn, TSGTKS được giả định tăng chậm hơn và đạt mức 115 vào năm 2020, sau đó trở về mức sinh học bình thường (105) vào năm 2030. Tác động khác nhau của các tình huống này sẽ được phản ánh trên cấu trúc giới tính của nhóm trẻ em sinh ra sau năm 2009.

Tình huống thứ ba giả thiết rằng TSGTKS chưa từng bao giờ tăng và luôn ở mức sinh học bình thường kể từ năm 1999. Vì sự gia tăng TSGTKS đã ảnh hưởng đến số liệu năm 2009, cơ cấu tuổi và giới năm 2009 được chỉnh theo cơ cấu dân số trẻ em của Thái Lan làm điểm tham chiếu, sau đó tiến hành dự báo số sinh trong 40 năm tiếp theo bằng cách sử dụng TSGTKS ở mức 105. Sự mô phỏng này chỉ ra cơ cấu tuổi của dân số Việt Nam sẽ như thế nào trong tình huống không có lựa chọn giới tính trước sinh.

PHỤ LỤC 9: MẤT CÂN BẰNG GIỚI TÍNH KHI SINH VÀ SỰ ƯA THÍCH CON TRAI

Mất cân bằng giới tính (được đo bằng TSGTKS) có liên quan chặt chẽ với mức độ ưa thích con trai của các cặp vợ chồng. Trên thực tế, TSGTKS có thể được hiểu như “sở thích đã được bộc lộ”, thể hiện rõ ràng sự mong muốn về giới tính của con cái trong các gia đình Việt Nam. Như đã trình bày ở trên, ưa thích con trai không phải là yếu tố duy nhất làm mất cân bằng TSGT. Các yếu tố khác như khả năng tiếp cận với công nghệ chọn lọc giới tính, tình trạng kinh tế-xã hội và mức sinh thấp cũng có ảnh hưởng đến tỷ lệ sinh trẻ em trai. Do vậy, đánh giá mức độ ưa thích con trai dựa trên số liệu nhân khẩu học vẫn là một vấn đề hóc búa vì các yếu tố khác cũng có thể có tác động đến mức độ TSGTKS quan sát được.

Tuy nhiên, số liệu TĐTDS 2009 của Việt Nam cho phép xây dựng các chỉ báo gián tiếp về tâm lý ưa thích con trai. Một trong những chỉ báo này là xu hướng sinh thêm con của các cặp vợ chồng trên cơ sở giới tính của các lần sinh trước. Sử dụng nhóm quần thể anh chị em ruột được xây dựng từ cơ cấu trẻ em trong gia đình, số liệu điều tra mẫu cho phép tính toán được tỷ trọng các gia đình sinh con thứ ba trong giai đoạn 1999-2009 trong tổng số các gia đình có ít nhất hai con vào năm 2009. Tỷ trọng này có tương quan tỷ lệ thuận với tỷ số lũy tiến theo thứ tự sinh hay xác suất có con thứ ba. Tỷ số lũy tiến này là rất quan trọng ở Việt Nam, khi mà mức sinh đã giảm từ ba xuống hai con trong vòng mười lăm năm qua. Tiếp theo, tỷ số giữa tỷ trọng và tỷ số lũy tiến nêu trên được tính toán theo tình trạng có hay không có con trai trong hai lần sinh trước. Tỷ số này sẽ xấp xỉ bằng 1 trong trường hợp không có sự ưa thích con trai. Không có gì ngạc nhiên khi kết quả thu được cho thấy xác suất tương đối này (sự khác biệt theo tỷ số lũy tiến theo thứ tự sinh) lớn hơn 1 ở tất cả các vùng lãnh thổ, chứng tỏ xác suất sinh con trai trong số các con thứ ba trong gia đình là rất cao. Tuy nhiên, khó có thể dự báo được mức độ biến thiên của tỷ số này theo các vùng lãnh thổ. Tỷ số này dao động từ 1,2 đến 5,7 qua các tỉnh, như trình bày trong Hình 24.

Một chỉ số khác phản ánh tâm lý ưa thích con trai trong số liệu TĐTDS là tần suất của các gia đình phụ hệ, được ước lượng bằng sự có mặt của con trai hay con gái đã kết hôn hiện đang sống cùng cha mẹ trong cùng một hộ gia đình. Mặc dù chế độ gia đình phụ hệ phổ biến ở Việt Nam, việc con gái sống cùng cha mẹ sau khi đã kết hôn cũng không phải là hiếm. Đây là một đặc điểm cho thấy hai hệ thống gia đình phụ hệ và mẫu hệ song song tồn tại ở một số vùng của Việt Nam, nhất là các tỉnh miền Trung và miền Nam. Chỉ số này được tính bằng tỷ lệ các hộ gia đình nhiều thế hệ trong đó chỉ có con trai đã lập gia đình sống cùng trong hộ. Phân tích này giới hạn trong các hộ gia đình có con đã lập gia đình và chủ hộ từ 40 tuổi trở lên. Tỷ lệ trung bình các hộ gia đình chỉ có con trai kết hôn sống trong hộ gia đình ở Việt Nam là 80,5%. Tỷ lệ cao như vậy không có nghĩa là hiếm gặp các hộ gia đình có con gái đã lập gia đình riêng nhưng vẫn sống cùng trong hộ gia đình. Hiện tượng con gái đã lập gia đình riêng nhưng vẫn sống cùng cha mẹ đẻ khá phổ biến ở nhiều tỉnh. Các ước lượng cấp tỉnh trong Biểu 24 minh họa sự khác biệt của hiện tượng này qua các vùng địa lý trong cả nước.

Hai chỉ số trên phản ánh các khía cạnh khác nhau của hệ thống xã hội ưa thích con trai. Chỉ số thứ nhất biểu hiện trực tiếp nỗ lực của các cặp vợ chồng trong việc đảm bảo có ít nhất một con trai, trong trường hợp cần thiết có thể có thêm con thứ ba sau khi đã sinh hai con gái. Như đã chỉ ra trên bản đồ, xu hướng sinh con thứ ba và đảm bảo lần sinh này là con trai thể hiện rõ nhất ở vùng Đồng bằng sông Hồng so với các vùng khác. Ngược lại, chỉ số thứ hai phản ánh hệ thống gia đình nhiều thế hệ tồn tại ở nhiều vùng rộng lớn trên cả nước, cho thấy các quan niệm về hệ thống gia đình phụ hệ khá đa dạng qua các vùng. Tương tự, chỉ số này cũng có xu hướng khác biệt đáng kể theo vùng. Khu vực phía Bắc Việt Nam có đặc trưng hệ thống gia đình phụ hệ mạnh mẽ hơn.

Các xu hướng trên hai Bản đồ nhìn chung là giống nhau. Các vùng tồn tại song song hai hệ thống gia đình tương ứng có mức TSGTKS theo thứ tự sinh thấp. Điều này gợi ý rằng chế độ gia đình phụ hệ và sự phụ thuộc vào con trai là một yếu tố quan trọng tác động đến tâm lý ưa thích con trai, được biểu hiện qua hành vi sinh sản. Kết quả phân tích cũng cho thấy hai bản đồ về ưa thích con trai trong Biểu 24 phù hợp với bản đồ trong Biểu 3 về TSGTTE. TSGTKS cao hơn ở những vùng có chế độ gia đình phụ hệ và nhu cầu có con trai cao hơn. Những xu hướng theo vùng diễn ra song song này gợi ý mặc dù các yếu tố cải thiện về cơ sở hạ tầng y tế, các điều kiện kinh tế xã hội và mức sinh thấp là có ảnh hưởng sâu sắc đến tình trạng mất cân bằng TSGTKS, thì ưa thích con trai vẫn là nhân tố đầu tiên và quan trọng nhất tác động đến lựa chọn giới tính.

Hình 24: Ước lượng hai chỉ số về sự ưa thích con trai theo tỉnh:

