
РЕГИОНАЛЬНОЕ РАЗНООБРАЗИЕ И РАСПРЕДЕЛЕНИЕ НАСЕЛЕНИЯ В ИНДИИ: ЭМПИРИЧЕСКИЕ ОЦЕНКИ И ТЕОРЕТИЧЕСКОЕ ОБОСНОВАНИЕ¹

Давыдов Денис Витальевич¹, д-р экон. наук, доц., вед. науч. сотр.
Шаповал Александр Борисович^{1,2}, д-р физ.-мат. наук, доц., проф.,
вед. науч. сотр.

¹ Лаборатория исследования социальных отношений и многообразия общества
Российской экономической школы, Сколковское шоссе, 45, каб. 2.46, Москва, Рос-
сия, 121353

² Национальный исследовательский университет, Высшая школа экономики, Мяс-
ницкая, 20, Москва, Россия, 101000; e-mail: ddavydov@nes.ru

Цель: развитие эмпирических методов и теоретических обоснований анализа неравномерности уровня региональной неоднородности на примере Индии. *Обсуждение:* в данной работе рассмотрены особенности региональной неоднородности населения по конфессиональному признаку на примере Индии. Построены количественные оценки индексов разнообразия населения на уровне регионов (штатов и отдельных территорий). Наблюдаемую кластеризацию полученных индексов нам удастся объяснить при помощи простой теоретико-вероятностной модели. *Результаты:* предложены способы эмпирического группирования регионов по уровню разнообразия, где в качестве индикатора диспропорций в уровне неоднородности используются различные индексы разнообразия. Предложена простая теоретическая модель, объясняющая структуру неравномерности уровня неоднородности для регионов Индии. Полученные эмпирические результаты и их теоретическое обоснование служат базой для дальнейших эмпирических исследований связи неоднородности и социально-экономического развития.

Ключевые слова: индексы неоднородности, региональное разнообразие, социально-экономическое развитие, Индия.

DOI: 10.17308/meps.2017.5/1646

Введение

Существующие обширные эмпирические исследования влияния неоднородности на экономическое развитие на уровне стран и макрорегионов обычно имеют характер «локальных успехов»: в отдельных случаях те или иные индексы позволяют получить статистически значимую связь между

¹ Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации, грант Правительства РФ (договор № 14.U04.31.0002).

степенью неоднородности населения по выделенному признаку и рассматриваемыми показателями социального и/или экономического развития (см., например, [18, 2, 7, 13]). Кроме того, большое внимание в научной литературе уделяется также выявлению и оценке влияния неоднородности обществ на миграционные процессы и возникновение конфликтов различного рода [19, 1, 5, 6, 12].

Более важную и детальную информацию содержат статистические данные на уровне регионов, однако использование региональных данных обычно является затруднительным в силу их неполноты, невысокой достоверности, отсутствия единой базы при межстрановом сопоставлении статистических показателей развития. Несмотря на это, подробное изучение отдельных факторов региональной неоднородности позволяет глубже исследовать феномен разнообразия в контексте экономического развития.

В данной работе мы уделяем основное внимание исследованию выявленных закономерностей в эмпирических данных, отражающих неоднородность населения Индии по конфессиональной принадлежности на уровне регионов, с использованием формальных свойств индексов неоднородности (см., в частности, [17, 4, 9, 10]). Построенные нами ранговые распределения для наиболее часто используемых индексов неоднородности позволяют на имеющихся эмпирических данных выявить достаточно устойчивую кластеризацию значений индексов. При этом степенной характер распределения на отдельных интервалах кластеризации может свидетельствовать об отсутствии характерных масштабов у изучаемой величины. С некоторыми оговорками здесь также можно усмотреть аналогию с идеями так называемого «закона Ципфа» [8, 14].

С другой стороны, мы показываем, что на используемых нами данных структура значений широко известных индексов неоднородности хорошо аппроксимируется гораздо менее распространенным «модальным» индексом, равным максимальной доле в соответствующих распределениях долей в каждом регионе страны. Ориентируясь далее на ранговое распределение указанного «модального» индекса, мы предлагаем простую вероятностную модель динамического распределения населения по регионам исследуемой страны. Модель относится к так называемым «урновым схемам», что характеризует ее как неэргодичную с точки зрения конечного распределения. Однако нас в данном случае интересует не стационарное (асимптотическое) распределение, а динамика рангового распределения максимальных долей. Мы демонстрируем, что получаемое нами модельное распределение максимальных долей, согласованное с распределением населения в целом по стране, позволяет достичь высокого уровня аппроксимации аналогичного рангового распределения, полученного на реальных данных.

Особенности конфессиональной неоднородности в Индии на уровне регионов

Индия является страной с высокой степенью неоднородности насе-

ния по всем базовым признакам, таким как владение языками, этническая или религиозная принадлежность. В некоторых регионах Индии язык, этнос и религия достаточно тесно связаны друг с другом, в иных случаях степень неоднородности по каждому из выделенных факторов существенно отличается на уровне отдельных штатов или более мелких территориальных единиц.

На уровне страны в целом Индия демонстрирует очень высокое языковое разнообразие. Среди более чем 400 языков и диалектов, распространенных на территории страны [16], Правительством Индии официально признаются 122 языка, из которых 22 имеют статус официальных (на общенациональном или региональном уровне). Наиболее распространенным языком в Индии, преимущественно в ее северной части, является хинди. Следующие наиболее часто используемые языки – это бенгали, телугу и марати. Более подробную информацию о структуре языкового разнообразия в Индии можно найти в базе данных Ethnologue [16]. Обзор и анализ особенностей распределения языков и проводимой в Индии языковой политики можно найти также в [11].

По религиозному составу населения Индия также отличается высокой сложностью и неоднородностью. Приведем некоторые наиболее показательные сведения, преимущественно основанные на анализе данных переписи населения Индии 2011 года ([3]; см. также Приложение 1). В целом на уровне страны около 80% населения исповедуют индуизм, представленный на сегодня множеством направлений, сект и религиозно-философских учений. В частности, в северной и центральной Индии широко распространены культы поклонения богу Вишну или одному из его воплощений – Кришне и Раме. В Тамилнаде, Карнатаке, а также некоторых районах Гималаев большинство населения поклоняются Шиве, а в Западной Бенгалии, Ассаме, Утта-Прадеш доминирующим является культ Шакти (матери-богини).

Индийская мусульманская община, составляющая около 14% населения Индии, сконцентрирована в штатах Джамму и Кашмир, Западной Бенгалии и западной части региона Утта-Прадеш. Она является одной из крупнейших в мире и превосходит по численности мусульманское население стран Ближнего Востока, уступая только общинам Индонезии, Пакистана и Бангладеш.

Христианство (менее 3% населения, преимущественно католики) сосредоточено на северо-востоке страны, в Бомбее и на юге Индии. Известны и другие религиозные общины, среди которых в том числе сикхи (Пенджаб и смежные районы), буддисты (Джамму и Кашмир, районы Гималаев), джайнисты, бахаисты, зороастрийцы (Бомбей, прибрежные районы Гуждарата), иудаисты (западное побережье), анимисты (северо-восток страны).

Описание исходных данных

Для формального количественного описания конфессиональной неоднородности на территории Индии в качестве исходных данных мы используем сведения из переписи населения Индии за 2011 г. в разрезе территориального

деления на уровне штатов [3]. Основные сведения приведены в Приложении 1.

Учитывая, что штат Телингана был сформирован в середине 2014 г., мы рассматриваем имеющиеся данные по двадцати восьми штатам и семи «объединенным территориям» Индии, отражающие конфессиональную принадлежность в разрезе следующих восьми групп: индуисты, мусульмане, христиане, сикхи, буддисты, джайнисты, представители прочих религий, атеисты. Соответствие используемых далее буквенных сокращений для обозначения штатов и территорий приведено в Приложении 1.

Формально, исходя из имеющихся данных, мы предполагаем, что Индия разделена на $N = 35$ условных «регионов», и в целом на территории страны можно выделить $n = 8$ перечисленных выше условных «конфессиональных групп», для упрощения называемых далее по тексту также «конфессиями» и «религиями». В каждом регионе j , $j = 1, \dots, N$, задан набор долей p_{ij} , $\sum_{i=1}^{n_j} p_{ij} = 1$, в которых представлены конфессии $i = 1, \dots, n_j$; где $n_j \leq n$ – это количество исповедуемых религий в регионе j . Другими словами, p_{ij} – это доля населения, исповедующего религию i в регионе j .

Пусть также $p_j = (p_{1j}, p_{2j}, \dots, p_{n_j j})$ – вектор соответствующих долей. Для каждого региона j на основании имеющихся данных мы вычисляем индексы разнообразия $A_j = A(p_j)$, $H_j = H(p_j)$, $R_j(a) = R(p_j, a)$, $a > 0, a \neq 1$, по следующим формулам:

$$A_j = 1 - \sum_{\tau=1}^{n_j} p_{\tau j}^2, \quad H_j = - \sum_{\tau=1}^{n_j} p_{\tau j} \ln p_{\tau j}, \quad R_j(a) = \frac{1}{1-a} \ln \left(\sum_{\tau=1}^{n_j} p_{\tau j}^a \right). \quad (1)$$

Индекс $A(\cdot)$ показывает вероятность того, что два наугад выбранных представителя региона принадлежат различным конфессиям. Индекс $H(\cdot)$ представляет стандартную энтропию. Индексы $R(\cdot, a)$ лежат в классе энтропийных индексов. При больших значениях параметра a их поведение характеризуется наибольшей долей p . При a , близких к нулю, индекс отражает количество различных долей (этносов), демонстрируя слабую зависимость от значений этих долей. При $a \rightarrow 1$ индексы $R_j(a)$ вырождаются в энтропию, которая отражает зависимость как от количества долей, так и от их конкретных значений. Подробнее о свойствах и особенностях данных индексов, а также о приложении энтропии к прогнозу редких событий см., например, [17, 15].

Особенности ранжирования регионов по индексам A , H и R

Вычислив по формулам (1) для каждого из рассматриваемых регионов Индии индексы A_j , H_j , $R_j(a)$ (при различных значениях параметра a , в том числе $a = 0.5$, $a = 2$), мы выявили, что порядок следования этих значений по перечню регионов в целом сохраняется, о чем свидетельствуют вычисленные коэффициенты парной корреляции Спирмена для соответствующих наборов индексов $A = \{A_j\}_{j=1}^W$, $H = \{H_j\}_{j=1}^W$, $R(a) = \{R_j(a)\}_{j=1}^W$, упорядоченных по совокупности регионов; в частности: $\rho(A, H) = 0.98$, $\rho(A, R(0.5)) = 0.90$.

При этом также выявляются следующие три характерные закономерности в отношении построенных наборов индексов и получаемых из них распределений.

Во-первых, можно говорить о группировании регионов по индексу A на интервалах его значений $[0.2, 0.22]$, $[0.28, 0.32]$, $[0.42, 0.51]$, что позволяет выделить три группы регионов. Тогда оставшиеся с обеих сторон распределения регионы, не вошедшие в указанные три группы, естественно также объединить, сформировав по совокупности пять групп регионов (рис. 1).

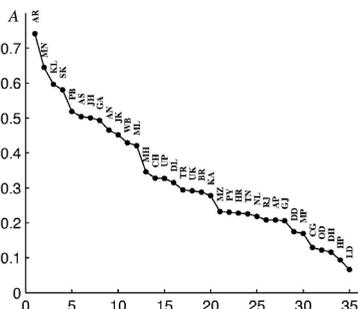


Рис. 1. Ранговое распределение индекса A по провинциям. Здесь и далее маркировка регионов, указанная на рисунках, соответствует кодам, приведенным в Приложении 1

Во-вторых, мы отмечаем, что при явном преобладании одной из долей, то есть при наличии выражено «доминирующей» конфессии в регионе, распределение индекса A мало отличается от распределения «модального» индекса M , отвечающего соответствующей наибольшей доле (рис. 2):

$$M_j = M(p_j) = 1 - p_{\max, j} = 1 - \max_{i=1, \dots, n} \{p_{ij}\}. \quad (2)$$

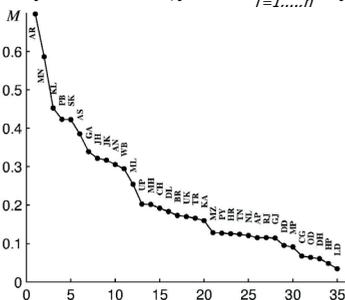


Рис. 2. Ранговое распределение индекса M по провинциям.

В-третьих, по совокупности анализируемых значений индексов $H_j, R_j(0.5), R_j(2)$ мы наблюдаем более крупное разбиение множества всех регионов на две группы с разными свойствами. Остановимся на последнем аспекте более подробно.

Разбиение регионов на две группы

Распределения индексов H и $R(a)$ в целом демонстрируют более гладкое изменение значений по сравнению с индексом A . Для них имеет место разделение регионов на две части по структуре распределения. Первые 10 точек (в левой части) рангового распределения H достаточно точно аппроксимируются степенной функцией (линейной в двойном логарифмическом масштабе, рис. 3). Последующие 19 точек также аккуратно аппроксимируются степенной функцией, но с другим наклоном.

У рангового распределения $R(0.5)$ разбиение на две группы возникает между 16-й и 17-й точками. Точность аппроксимаций у этого распределения выше, и точка перехода становится более явной (рис. 4).

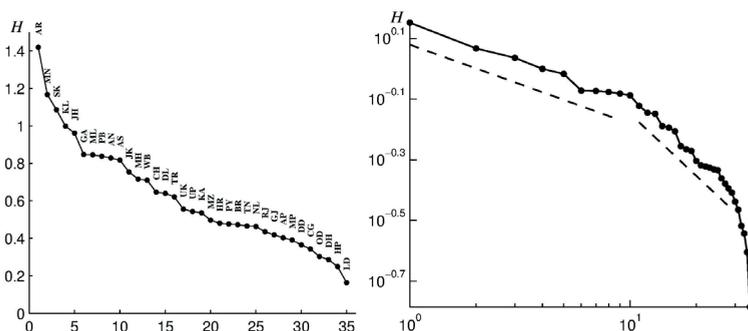


Рис. 3. Ранговое распределение индекса H по провинциям в линейном (слева) и двойном логарифмическом (справа) масштабах. Линейные аппроксимации имеют наклоны -0.26 и -0.68

Мы также проверили (но не приводим здесь дополнительных графических иллюстраций), что разбиение распределений $R(a)$ на две части, из которых правая часть близка к линейной, характерна для всего семейства распределений $R(a)$ (экстремальные значения параметра a , $a > 5$ и $a < 0.1$ не рассматривались при численном анализе, поскольку индексы в этих случаях близки к другим известным и более простым характеристикам). Например, при $a = 2$ ранговое распределение индекса $R(a)$ сохраняет квазистепенной участок в левой части, тогда как в правой части начинают образовываться кластеры регионов, которые более отчетливо видны на распределениях индекса A (рис. 5).

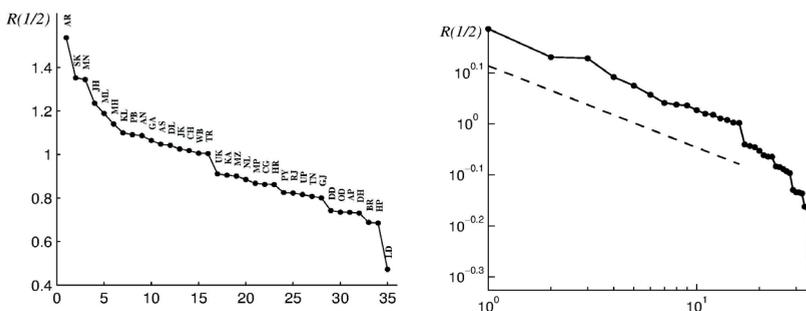


Рис. 4. Ранговое распределение индекса $R(0.5)$ по провинциям в линейном (слева) и двойном логарифмическом (справа) масштабах. Аппроксимация имеет наклон -0.16 .

Выявленное нами свойство группирования значений индексов разнообразия для разных классов индексов подтверждается также дополнительным анализом географического распределения регионов, попавших в соответствующие группы. В частности, мы сопоставили значения индекса $R(0.5)$ с географическим положением соответствующих регионов (рис. 6). Центральная часть Индии характеризуется умеренными (средними) значениями индекса, которые возрастают по мере приближения к границам государства.

При заданном шаге в 0.1 по синим участкам карты можно перемещаться, изменяя значения индекса не более чем на 0.1. Провинции, которые расположены на востоке страны, как правило, имеют более высокие значения индекса, достигая максимума на границе с Китаем.

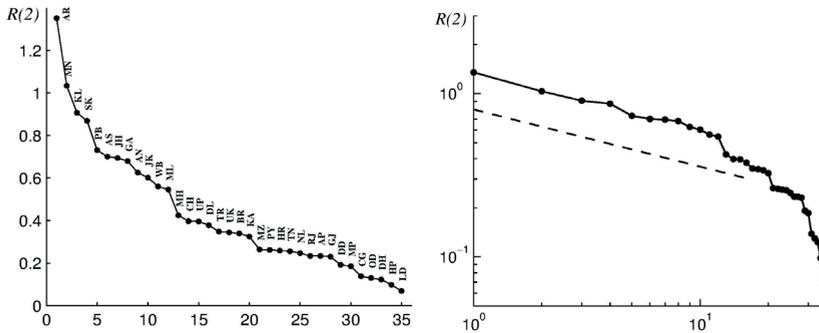


Рис. 5. Ранговое распределение индекса $R(2)$ по провинциям в линейном (слева) и двойном логарифмическом (справа) масштабах. Аппроксимация имеет наклон -0.35

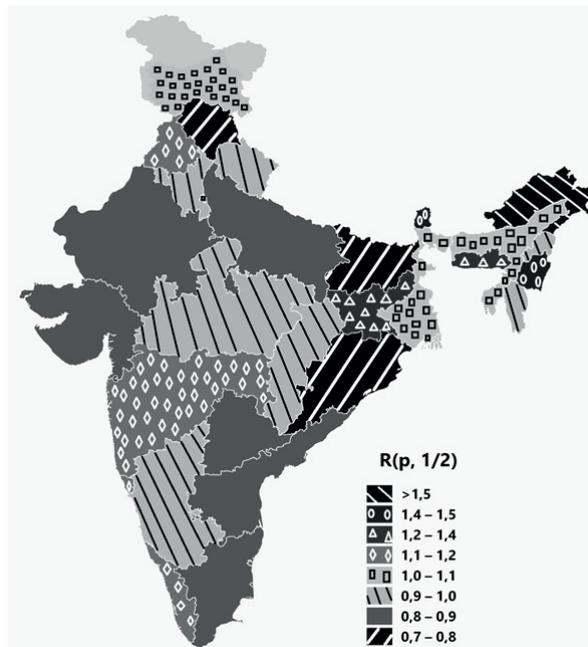


Рис. 6. Регионы Индии, упорядоченные в соответствии с индексом $R(0.5)$

Напомним, что степень неоднородности региона увеличивается со значением индекса $R(a)$, при $a < 1$. Поэтому большинство территорий страны, включая столицу Нью-Дели, характеризуются умеренной неоднородностью, а ее изменение от центра страны на восток происходит негладко. Уровень неоднородности уменьшается по направлению к Непалу и к побережью, но увеличивается по направлению к Республике Бангладеш. Изменения уровня

неоднородности при переходе из региона в регион в восточной части страны выглядят беспорядочными. По сравнению с остальной частью страны подобное отсутствие простой закономерности одновременно накладывается на высокое среднее значение индексов неоднородности в этой части Индии.

Опираясь на все приведенные выше результаты количественного анализа индексов неоднородности, далее мы предлагаем рассмотреть простую вероятностную модель расселения людей по регионам страны, позволяющую объяснить и дополнить с теоретической точки зрения выявленные нами эмпирические закономерности.

Модель

Построенное нами по всем регионам Индии ранговое распределение долей наиболее представительных конфессий в каждом из регионов для части из них хорошо согласуется со следующей простой моделью расселения.

Будем последовательно расселять k жителей, представляющих n конфессий, по N регионам, где $n = 8$ и $N = 35$ совпадают с наблюдаемыми величинами для рассматриваемых нами данных, а k – используемый нами масштабирующий параметр модели.

Пусть для определенности очередной житель исповедует религию i . Выберем случайным образом регион из нашего списка. В этом регионе j , $j = 1, \dots, N$, уже проживает некоторое количество жителей, исповедующих ту же религию i . Через p_{ij} обозначим их долю среди всех жителей региона j . Доли p_{ij} изменяются в процессе расселения. Мы определим, что очередной житель направляется в регион j с вероятностью p_{ij} , имеющейся на момент расселения данного жителя. В противном случае, с вероятностью $(1 - p_{ij})$, процедура выбора региона, описанная выше, а именно, (i) выбор региона наугад среди всех регионов, (ii) решение жителя в нем остаться с вероятностью p_{ij} , где j – выбранный регион, продолжается до тех пор, пока регион не будет выбран.

Чтобы описанная процедура функционировала, требуется определить начальные условия для вероятностей p_{ij} , $i = 1, \dots, n$, $j = 1, \dots, N$. Для определенности положим начальные значения $p_{ij} = 1/n$ для всех индексов i и j , то есть будем считать, что в каждом регионе до начала процедуры расселения k (новых) человек представительство всех конфессий равномерно (равновероятно при случайной встрече с представителем любой из них). Непосредственные вычисления показывают, что в результате применений правил (i) и (ii) представитель конфессии i отправится в регион j с вероят-

ностью $\tilde{p}_{ij} = \frac{p_{ij}}{\sum_{j=1}^n p_{ij}}$, которая вслед за p_{ij} также меняется по мере расселения

жителей по регионам. Именно этой формулой естественно пользоваться далее при компьютерном моделировании процессов расселения на основании введенных правил (i) и (ii).

Расселив жителей страны по указанным правилам, определим в каждом регионе наибольшую долю: $p_{max,j} = \max_{i=1,\dots,n} p_{ij}$.

Ранговое распределение наибольшей доли, разумеется, зависит от каждой конкретной (случайной) реализации модельного механизма. Для исключения влияния случайной компоненты мы проводим соответствующую процедуру осреднения. А именно, зафиксировав количество k расселяемых жителей, мы запускаем $v = 100$ численных реализаций модели, в каждой из которых вычисляем максимальную долю $p_{max,j}^v$ для всех $N = 35$ рассматриваемых регионов Индии, затем формируем из них множество $P^v = \{p_{max,j}^v\}_{j=1}^{35}$ и сортируем его в порядке возрастания. С целью упрощения мы также используем далее обозначение $\{p_{max,j}^\mu\}_{j=1}^{35}$ для получившегося (упорядоченного) множества. Для каждого j (в сортированном множестве) вычисляется среднее и стандартное отклонение вычисленных долей:

$$\bar{p}_{max,j} = \frac{1}{v} \sum_{\mu=1}^v p_{max,j}^\mu, \quad \sigma_j = \sqrt{\frac{1}{v} \sum_{\mu=1}^v (p_{max,j}^\mu - \bar{p}_{max,j})^2}.$$

Вновь упрощая обозначения, возвращаемся от $\bar{p}_{max,j}$ к $p_{max,j}$. На рис. 7 показаны (осредненные) ранговые распределения вычисленного «модального» индекса M (см. выше формулу (2)), полученные при расселении k человек, где k последовательно принимает значения, равные $2, 4, 8, 16 \times 10^4$.

Поточечная сходимость фрагмента правой части представленных на рис. 7 (слева) графиков, демонстрируется на рис. 7 (справа). При последовательном удвоении количества жителей (параметра k) графики монотонно смещаются вверх, однако амплитуда такого смещения уменьшается (рис. 7, справа). Аналогичное уменьшение сдвига мы наблюдаем и на других частях графика, но отдельно не иллюстрируем.

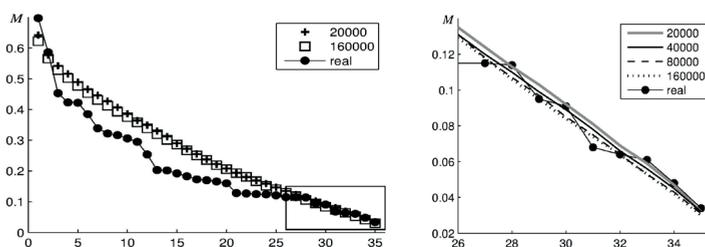


Рис. 7. Ранговое распределение максимальных долей $p_{max,j}$ по регионам $j = 1, \dots, 35$; полное – слева, фрагмент, помещенный в рамку, – справа; $k = 2, 4, 8, 16 \times 10^4$

Стандартное отклонение $\{\sigma_j\}_{j=1}^{35}$, как функция от j , стабилизируется с увеличением k при достижении этим параметром значения 80000 (рис. 8).

В результате ранговое распределение «коридора значений» $\{1 - p_{max,j} \pm \sigma_j\}_{j=1}^{35}$ уже при $k = 160000$ позволяет получить представление о предельном модельном распределении «модального» индекса M (рис. 9).

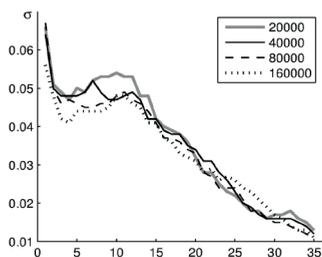


Рис. 8. Стандартное отклонение σ_j максимальных долей $p_{max,j}$ по регионам $j=1, \dots, 35$, вычисленное по каждому региону

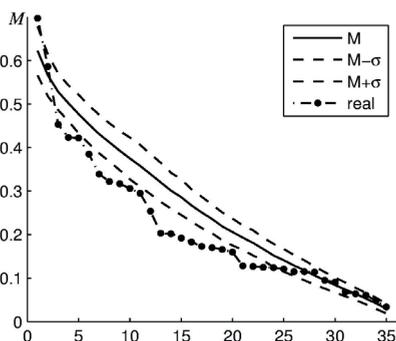


Рис. 9. Ранговое распределение $\{1 - p_{max,j} \pm \sigma_j\}_{j=1}^{35}$ при $k=16000$ максимальных долей $p_{max,j}$ (сплошная линия) со стандартной ошибкой (пунктир), равной среднеквадратичному отклонению, по регионам $j=1, \dots, 35$

Как мы видим на рис. 9, модельное распределение достаточно точно приближает наблюдаемое ранговое распределение в области больших значений максимальной доли. В эту область входят 9 точек из 36, то есть граница области является верхним квартилем. Полученный результат дополняет описанную выше эмпирическую закономерность по группированию регионов по степени конфессиональной неоднородности. А именно, при группировании регионов по близким значениям для индекса А, указанные 9 точек были объединены в группу по остаточному признаку, поскольку соответствующие им значения индекса были меньше, чем поддающиеся группировке значения.

Таким образом, мы полагаем, что представительство различных религий в регионах с доминирующим количеством индивидов из одной (преобладающей) религии формируется в соответствии с изложенными здесь принципами элементарной модели (правилами (i) и (ii)), которые в практическом контексте означают, во-первых, случайный выбор места расселения (при прочих равных) и, во-вторых, увеличение относительной численности соответствующей модальной конфессии. Напротив, группирование других регионов по степени неоднородности, где доля доминирующей религии не столь велика, скорее объясняется отклонением от указанных элементарных процессов.

В рамках предложенной нами модели более высокие значения $P_{max,j}$ в основной части графика (регионы 5-25), что соответствует «провалу» реальных данных по сравнению с модельными значениями (рис. 9), достигаются, если вероятность поселиться среди представителей «своей» конфессии в этих регионах увеличивается за счет уменьшения вероятности оказаться в регионах 1-4, где наблюдаются наиболее высокие значения уровня разнообразия (это результат проведенного нами численного моделирования, не подтверждаемый здесь дополнительными иллюстрациями). Соответствующее (небольшое) перераспределение вероятностей индуцирует наблюдаемое ранговое распределение максимальной доли, у которого средняя часть распределения (регионы 5-25) «поднимется» за его счет левой части (регионы 1-4). Такое поведение может быть, в частности, связано с миграцией в регионы, где близких по религии индивидов «достаточно много, но не слишком много».

Заключение

В данной работе мы сочетаем простое теоретическое обоснование наблюдаемой на практике неравномерной неоднородности на уровне регионов различных стран в виде вероятностной модели расселения с эмпирическим анализом структуры такой неравномерной неоднородности. В качестве примера мы используем данные по конфессиональной неоднородности на уровне штатов и обособленных территорий Индии – одной из наиболее неоднородных по основным критериям стран мира.

Мы предлагаем использовать группирование регионов по уровню разнообразия в качестве индикатора различий в уровне неоднородности на основе различных индексов, что само по себе может служить дополнительной структурной информацией в последующих эконометрических тестах при оценке влияния неоднородности на социальное и экономическое развитие.

Отдельно подчеркнем, что предлагаемая нами модель может объяснить, и что она объяснить не может. Модель описывает эволюцию общества от равномерного распределения представителей различных конфессий в соответствии с заданными нами естественными и простыми правилами (*i*) и (*ii*). Оказывается, что возникающее распределение «модальной» доли (доли населения доминирующей конфессии в данном регионе) в целом хорошо согласуется с наблюдениями. Однако все распределение всех долей в целом по регионам страны, полученное нами в результате моделирования, с имеющимися наблюдениями согласуется лишь частично, а потому при необходимости получения модельных представлений обо всей структуре расселения требует более тонкого моделирования, которое находится за пределами цели данной работы.

Приложение 1

Регионы Индии, их условные обозначения и данные о религиозной принадлежности населения в региональном разрезе (тыс. чел.)

№ п/п	Название штата	Код штата	Население (всего)	Религиозная принадлежность							
				Индуисты	Мусульмане	Христиане	Сикхи	Буддисты	Джайанисты	Другие	Атеисты
1	Uttar Pradesh	UP	199 812	159 313	38 484	356	644	206	213	14	583
2	Maharashtra	MH	112 374	89 703	12 971	1 080	223	6 531	1 400	179	286
3	Bihar	BR	104 099	86 079	17 558	129	24	25	19	13	252
4	West Bengal	WB	91 276	64 386	24 655	659	64	283	60	942	228
5	Andhra Pradesh	AP	84 581	74 824	8 082	1 130	40	37	54	10	404
6	Madhya Pradesh	MP	72 627	66 007	4 775	213	151	216	567	600	98
7	Tamil Nadu	TN	72 147	63 188	4 229	4 418	15	11	89	7	189
8	Rajasthan	RJ	68 548	60 657	6 215	96	873	12	622	5	68
9	Karnataka	KA	61 095	51 317	7 893	1 143	29	96	440	11	166
10	Gujarat	GJ	60 440	53 534	5 847	316	58	30	580	16	58
11	Odisha	OD	41 974	39 300	912	1 162	22	14	9	478	77
12	Kerala	KL	33 406	18 282	8 873	6 141	4	5	4	8	88
13	Jharkhand	JH	32 988	22 376	4 794	1 419	71	9	15	4 236	68
14	Assam	AS	31 206	19 181	10 679	1 166	21	55	26	27	51
15	Punjab	PB	27 743	10 678	535	348	16005	33	45	11	88
16	Chhattisgarh	CG	25 545	23 820	515	491	70	70	62	495	23
17	Haryana	HR	25 351	22 171	1 781	50	1 244	8	53	3	42
18	Jammu & Kashmir	JK	12 541	3 567	8 567	36	235	113	2	2	20
19	Uttarakhand	UK	10 086	8 369	1 407	38	236	15	9	1	12
20	Himachal Pradesh	HP	6 865	6 533	150	13	80	79	2	1	8
21	Tripura	TR	3 674	3 064	316	160	1	125	1	2	5
22	Meghalaya	ML	2 967	342	130	2 213	3	10	1	258	10
23	Manipur	MN	2 856	1 182	240	1 179	2	7	2	234	11
24	Nagaland	NL	1 979	173	49	1 740	2	7	3	3	2
25	Goa	GA	1 459	964	122	366	1	1	1	0	3
26	Arunachal Pradesh	AR	1 384	402	27	419	3	163	1	363	7
27	Mizoram	MZ	1 097	30	15	956	0	93	0	1	1
28	Sikkim	SK	611	353	10	61	2	167	0	16	2
29	Chandigarh	CH	1 055	853	51	9	138	1	2	0	1
30	Puducherry	PY	1 248	1 089	76	79	0	0	1	0	2
31	Andaman & Nicobar Islands	AN	381	264	32	81	1	0	0	1	1
32	Dadra & Nagar Haveli	DH	344	323	13	5	0	1	1	0	0
33	Daman & Diu	DD	243	220	19	3	0	0	0	0	0
34	Lakshadweep	LD	64	2	62	0	0	0	0	0	0
35	Nct Of Delhi	DL	16 788	13 712	2 159	146	571	18	166	2	14
Итого по стране			1 210 855	966 257	172 245	27 820	20833	8 443	4 452	7 938	2 867

*) Штат Телингана сформирован в 2014 году и не включен в данный перечень

Источник: [3]

СПИСОК ИСТОЧНИКОВ

1. Adsera A., Pytlikova M. *Language and Migrations* / In: V. Ginsburgh, S. Weber (Eds.) *The Palgrave Handbook of Economics and Language*. Palgrave Macmillan UK, 2016. pp. 342-372.
2. Alesina A., Devleeschouwer A., Easterly W., Kurlat S., Wacziarg R. Fractionalization // *Journal of Economic Growth*, 2003, Vol. 8, no. 2, pp. 155-194.
3. Census-2011 (India): The official statistics report. Available at: <http://www.censusindia.gov.in> (accessed: 12.03.17).
4. Davydov D., Weber S. A simple characterization of the family of diversity indices // *Economics Letters*, 2016, vol. 147, pp. 121-123.
5. Esteban J., Mayoral L., Ray D. Ethnicity and Conflict: An Empirical Study // *American Economic Review*. 2012, Vol. 102, no. 4, pp. 1310-1342.
6. Fearon J.D., Laitin D.D. Ethnicity, Insurgency, and Civil War // *American Political Science Review*, 2003, Vol. 97, no. 1, pp. 75-90.
7. Fearon J. Ethnic and Cultural Diversity by Country // *Journal of Economic Growth*, 2003, Vol. 8, no. 2, pp. 195-222.
8. Gabaix X. Zipf's Law for Cities: An Explanation // *Quarterly Journal of Economics*, 1999, Vol. 114, no. 3, pp. 739-776.
9. Ginsburgh V., Weber S. *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton, NJ: Princeton University Press, 2011.
10. Greenberg J. The measurement of linguistic diversity // *Language*, 1956, Vol. 32, pp. 109-115.
11. Laitin D. Language policy and political strategy in India // *Policy Sciences*, 1989, Vol. 22, pp. 415-436.
12. Montalvo J. G., Reynal-Querol M. Ethnic Polarization, Potential Conflict and Civil Wars // *American Economic Review*, 2005, Vol. 95, no. 3, pp. 796-816.
13. Montalvo J.G., Reynal-Querol M. Religious Polarization and Economic Development // *Economics Letters*, 2003, Vol. 80, no. 2, pp. 201-210.
14. Shapoval A.B. and M. Shnirman The BTW mechanism on a self-similar image of a square: A path to unexpected exponents // *Physica A: Statistical Mechanics and its applications*, 2012, Vol. 391, no. 1-2, pp. 15-20.
15. Shapoval A.B., Shnirman M.G. Strong Events in the Sand-Pile Model // *International Journal of Modern Physics C*, 2004, Vol. 15, no. 2, pp. 279-288.
16. Simons G.F., Fennig C.D. (Eds.). *Ethnologue: Languages of the World. Twentieth edition*. Dallas, TX: SIL International, 2017. Available at: <http://www.ethnologue.com> (accessed: 12.03.17).
17. Акчурина Д.Д., Вебер Ш., Давыдов Д.В., Крутиков Д.В., Хазанов А.А. Измерение разнообразия: теория и социально-экономические приложения // *Современная экономика: проблемы и решения*, 2015, no. 2, с. 8-28.
18. Вебер Ш., Давыдов Д.В. Экономика разнообразия: подходы, методы, результаты // *Экономика и математические методы*, 2015, т. 51, no. 4, с. 3-13.
19. Вебер Ш., Давыдов Д.В. Этнокультурные факторы, социальные процессы, конфликты и экономическое развитие // *Современная экономика: проблемы и решения*, 2015, no. 8, с. 129-143.

REGIONAL DIVERSITY AND POPULATION DISTRIBUTION IN INDIA: EMPIRICS AND THEORETICAL APPROACH

Davydov Denis Vitalievich¹, Dr. Sc. (Econ.), Assoc. Prof., research fellow at NES CSDSI

Shapoval Alexander Borisovich^{1,2}, Dr. Sc. (Math.), Assoc. Prof., Prof. at NRU HSE, research fellow at NES CSDSI

1 New Economic School Center for the Study of Diversity and Social Interactions, Skolkovskoe shosse, 45, of. 2.46, Moscow, Russia, 121353

2 National Research University Higher School of Economics, Myasnitskaya st., 20, Moscow, Russia, 101000; e-mail: ddavydov@nes.ru

Purpose: development of empirical methods and theoretical foundations for the analysis of the variations of regional population heterogeneity in India.

Discussion: based on different quantitative estimations of diversity, we pay the main attention to the problem of regional confessional heterogeneity disparities in India. Calculating and comparing several diversity indices of regional heterogeneity we reveal the property of grouping the regions in some clusters which demonstrate rather different properties. The latter, in turn, can be explained with a simple probabilistic model of population distribution. *Result:* we introduce and examine a quantitative computational model that matches the revealed empirical evidence of rank distributions for different diversity indices which measure confessional heterogeneity of the states and territories of India. The obtained results are expected to provide a new foundation for further study of heterogeneity in the context of social and economic development.

Keywords: heterogeneity indices, regional diversity, social and economic development, India.

References

1. Adsera A., Pytlikova M. *Language and Migrations* / In: V. Ginsburgh, S. Weber (Eds.) *The Palgrave Handbook of Economics and Language*. Palgrave Macmillan UK, 2016, pp. 342-372.
2. Alesina A., Devleeschouwer A., Easterly W., Kurlat S., Wacziarg R. Fractionalization. *Journal of Economic Growth*, 2003, vol. 8, no. 2, pp. 155-194.
3. Census-2011 (India): The official statistics report. Available at: <http://www.censusindia.gov.in> (accessed: 12.03.17)
4. Davydov D., Weber S. A simple characterization of the family of diversity indices. *Economics Letters*, 2016, vol. 147, pp. 121-123.
5. Esteban J., Mayoral L., Ray D. Ethnicity and Conflict: An Empirical Study. *American Economic Review*, 2012, vol. 102, no. 4, pp. 1310-1342.
6. Fearon J.D., Laitin D.D. Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 2003, vol. 97, no. 1, pp. 75-90.
7. Fearon J. Ethnic and Cultural Diversity by Country. *Journal of Economic Growth*, 2003, vol. 8, no. 2, pp. 195-222.
8. Gabaix X. Zipf's Law for Cities: An

- Explanation. *Quarterly Journal of Economics*, 1999, vol. 114, no. 3, pp. 739-776.
9. Ginsburgh V., Weber S. *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton, NJ: Princeton University Press, 2011.
10. Greenberg J. The measurement of linguistic diversity. *Language*, 1956, vol. 32, pp. 109-115.
11. Laitin D. Language policy and political strategy in India. *Policy Sciences*, 1989, vol. 22, pp. 415-436.
12. Montalvo J.G., Reynal-Querol M. Ethnic Polarization, Potential Conflict and Civil Wars. *American Economic Review*, 2005, vol. 95, no. 3, pp. 796-816.
13. Montalvo J.G., Reynal-Querol M. Religious Polarization and Economic Development. *Economics Letters*, 2003, vol. 80, no. 2, pp. 201-210.
14. Shapoval A.B. and M. Shnirman The BTW mechanism on a self-similar image of a square: A path to unexpected exponents. *Physica A: Statistical Mechanics and its applications*, 2012, vol. 391, no. 1-2, pp. 15-20.
15. Shapoval A.B., Shnirman M.G. Strong Events in the Sand-Pile Model. *International Journal of Modern Physics C*, 2004, vol. 15, no. 2, pp. 279-288.
16. Simons G.F., Fennig C.D. (Eds.). *Ethnologue: Languages of the World. Twentieth edition*. Dallas, TX: SIL International, 2017. Available at: <http://www.ethnologue.com>. (accessed: 12.03.17).
17. Akchurina D.D., Veber Sh., Davydov D.V., Krutikov D.V., Khazanov A.A. Izmerenie raznoobraziia: teoriiia i sotsial'no-ekonomicheskie prilozheniia. *Sovremennaia ekonomika: problemy i resheniia*, 2015, no. 2, pp. 8-28. (In Russ.)
18. Veber Sh., Davydov D.V. *Ekonomika raznoobraziia: podkhody, metody, rezul'taty. Ekonomika i matematicheskie metody*, 2015, t. 51, no. 4, pp. 3-13. (In Russ.)
19. Veber Sh., Davydov D.V. Etnokul'turnye faktory, sotsial'nye protsessy, konflikty i ekonomicheskoe razvitie. *Sovremennaia ekonomika: problemy i resheniia*, 2015, no. 8, pp. 129-143. (In Russ.)