

## Моделирование межгруппового индекса цен покупок продуктов питания в зависимости от микроэкономических характеристик домашних хозяйств

Матыцин М.С.

В работе рассматривается вопрос об эндогенности цен покупок и их зависимости от различных микроэкономических характеристик домохозяйств. Соответствующая гипотеза находит свое подтверждение при моделировании межгрупповой динамики индекса цен покупок продуктов питания по группам домохозяйств в зависимости от суммы подушевых расходов на питание. Такое моделирование требует наличия соответствующей информации в данных выборочных обследований, доступной только в RLMS. Также предлагается механизм совместного использования информации RLMS и Выборочного обследования бюджетов домашних хозяйств Росстата (ВОБДХ), в том числе для учета неоднородности цен.

**Ключевые слова:** домашние хозяйства, потребление, эндогенность цен, межгрупповые индексы цен, выборочные обследования домохозяйств.

### 1. Введение

Мы ставили своей задачей показать, что в условиях, когда рассмотрение домашних хозяйств как единого агента, не учитывающее его структуру, может существенным образом повлиять на точность получаемых прогнозов потребления, необходима статистическая проверка гипотезы наличия неоднородности этого агента. В случае ее принятия моделирование должно вестись с учетом такой неоднородности. Для корректного построения моделей необходимо выяснить вопрос о наличии требуемых статистических данных и в случае необходимости интегрировать информацию из разных источников.

Основной целью работы является проверка гипотезы о наличии эндогенности цен, т.е. о реализации потребителем возможности выбирать цены покупок отдельных

---

Автор выражает искреннюю благодарность за помощь и научное руководство Ершову Э.Б.

**Матыцин М.С.** – аспирант кафедры математической экономики и эконометрики факультета экономики Государственного университета – Высшей школы экономики, инженер Лаборатории макроструктурного моделирования экономики России ГУ ВШЭ. E-mail: matm@yandex.ru

Статья поступила в Редакцию в августе 2010 г.

товаров в процессе реализации своего спроса. Для проверки такой гипотезы и предложения механизма учета неоднородности цен покупок последовательно в соответствующих разделах статьи решаются несколько задач.

Во втором разделе поясняется возможность возникновения эндогенности цен и формулируется гипотеза о зависимости межгруппового индекса цен от микроэкономических характеристик домохозяйств, в том числе от уровня дохода.

Третий раздел посвящен описанию доступных источников статистической информации, необходимых для проверки такой гипотезы. В российской статистической практике в этих целях могут быть использованы два выборочных бюджетных обследования домохозяйств – ВОБДХ, проводимое Росстатом, и RLMS, различающиеся по структуре и методу сбора данных. Из них только RLMS содержит информацию об индивидуальных ценах покупок, требуемую для проверки выдвинутой гипотезы. В этом разделе проверяется гипотеза эндогенности цен в форме наличия зависимости межгрупповой динамики индекса цен покупок продуктов питания от суммы расходов на питание на человека. Такая гипотеза не отвергается при различном выборе числа групп и способа распределения по ним домохозяйств.

В четвертом разделе предлагается механизм совместного использования информации из двух выборок – ВОБДХ и RLMS – для моделирования структуры расходов домашних хозяйств с учетом неоднородности цен покупок в зависимости от микроэкономических характеристик домохозяйств.

В заключении приводятся выводы.

## 2. Гипотеза эндогенности цен

В научной и учебной литературе существуют несколько принципиально различающихся подходов к анализу и моделированию поведения потребителей, под которыми обычно понимаются отдельные индивидуумы или домашние хозяйства. Наиболее известная и продвинутая версия теории такого поведения базируется на предположении, согласно которому оно соответствует максимизации потребителем своей функции полезности на его бюджетном ограничении при экзогенно заданных для потребителя ценах и уже определенных общих потребительских расходах. Последние определяются в рамках модельных построений, характеризующих динамику расходов, ожидаемых доходов и цен, сбережений, заимствований и возврата долгов домашнего хозяйства.

Предположение об экзогенности цен на покупаемые товары и услуги вполне оправдано микроэкономическим и статическим рассмотрением проблемы, когда индивидуальный потребитель не оказывает влияния своим спросом на цены. Однако на практике потребление может быть устроено иначе. Потребитель, предъявляя спрос на тот или иной товар или услугу из некоторой группы, одновременно с выбором требуемого количества товара и качества (категории, торговой марки и т.д.), во многих случаях имеет возможность выбрать конкретный товар из некоторой группы товаров (услуг) и цену, по которой он покупается. Выбор товара с не минимальной ценой не вступает в прямое противоречие с предположением о рациональности индивида, может быть объяснен целым рядом факторов, подробное рассмотрение которых, как правило, опускается при описании классических микроэкономических моделей. Так, некоторые категории потребителей готовы платить более высокую цену за товары (пусть даже и того же качества), желая, например, сэкономить время покупки (т.е.

приобретая в ближайшем магазине, где цены могут отличаться) или получить более высокий уровень сервиса (покупая в магазине более высокого класса). Другими мотивами могут быть желание избежать дополнительных транзакционных издержек по поиску наилучшей цены или из соображений учета требований коллективного характера использования товара. Результат такого выбора цен покупок будем называть эндогенными ценами, т.е. определяемыми в процессе покупки, в отличие от традиционного подхода, имеющего дело с экзогенно задаваемыми ценами.

Для коротких периодов экзогенность цен, по-видимому, оправдывается и при рассмотрении поведения групп потребителей или их совокупности в масштабах экономики страны или ее региона. Однако такая экзогенность средних цен для группы товаров не эквивалентна, по нашему мнению, их равенству для различных, предполагаемых однородными групп потребителей. То есть разные категории (классы) потребителей внутри своей группы могут не оказывать влияния на цены конкретных товаров и услуг. Сами цены покупок, которые они реализуют в результате своего выбора, могут достаточно сильно различаться в зависимости от целого ряда характеристик этой категории потребителей – уровня их доходов или расходов, региона проживания, типа населенного пункта и т.д.

В условиях, когда цены покупок на группы товаров и услуг для разных групп потребителей существенно различаются и в динамике изменяются веса численностей таких групп и их доходов в их общей совокупности, моделирование макроэкономического потребительского поведения совокупности домашних хозяйств как поведения одного представительного агента может оказаться не учитывающим существенные и динамичные стороны изучаемого объекта. Поэтому должна быть подтверждена или опровергнута гипотеза об эндогенности (или экзогенности) рыночных цен, с которыми сталкиваются домашние хозяйства. Эту гипотезу необходимо корректно сформулировать. В зависимости от того, оправдывается ли она или будет отвергнута, можно будет уточнять и развивать постановку проблемы макроэкономического и макроструктурного моделирования и прогнозирования потребительских расходов домашних хозяйств.

Проблема эндогенности цен тесно связана с проблемой выбора потребителем качества продукта. Традиционно во многих моделях цена рассматривается как сигнал качества продукта [12, 16]. Таким образом, выбор цен покупок для разных категорий потребителей может быть свидетельством различного предпочтения качества, на которое они ориентируются по ценам. Но более высокие цены покупок для некоторых групп домохозяйств могут быть вызваны в том числе разным качеством процесса покупки, а не только самого товара (например, расположением магазина, дополнительными услугами, включенными в цены товаров: парковкой и т.д.).

Экономическая теория не конкретизирует то, для какого набора товаров и услуг она описывает потребительское поведение одного домашнего хозяйства или некоторой их совокупности. Такая конкретизация делается в исследованиях эмпирической направленности, использующих статистические данные. Структура последних относительно консервативна и накладывает ограничения на область практического использования соответствующей теории. Из общих соображений очевидно, что множество товаров и услуг, на котором конкретное домашнее хозяйство способно формировать и реализовать свои предпочтения, не является универсальным, зависит от многих его характеристик и изменяется со временем. Для отдельного домашнего хозяйства анализ потребительского поведения неизбежно сводится к вербальному его

описанию. Возможности моделирования такого поведения возникают при рассмотрении группы потребителей, предполагаемой однородной. Целесообразности выделения таких групп и применения к ним теории, использующей функции полезности, посвящена Глава 18 «Руководства по индексу потребительских цен: Теория и практика», подготовленного рядом международных организаций [11]. Но применение такого подхода наталкивается на серьезные практические трудности.

В исследованиях, использующих российские данные, отмечаются и характеризуются различия цен, по которым потребители приобретают товары из одной группы. Так, Варшавский А.Е. в работе, посвященной исследованию внутреннего рынка продуктов питания [2], приходит к заключению о сегментации рынка пищевых продуктов, характеризуемой двумя тенденциями: «рост цен и качества – для богатых, снижение качества при постоянной цене – для бедных». Очевидно, что эти тенденции проявляются в дифференциации средних цен покупок для потребителей с разными доходами.

Различие цен, по которым приобретаются товары и услуги различными группами домашних хозяйств, было осознано и в теоретическом плане. Петер Калман в двух своих работах [15, 16] включил в число аргументов функции полезности не только приобретаемые или потребляемые количества ( $q$ ) товаров и услуг, но их цены ( $p$ ). Такое включение Калман мотивировал, ссылаясь на работы Маршака [17], Веблена [19] и других экономистов, в которых цены покупок связывались с качеством продуктов. Калман показал, что такое обобщение понятия «функция полезности» приводит к возможности получения неоднородных по ценам и суммарным расходам функций спроса. Но в этих работах идея не была развита, не доведена до выбора класса соответствующих функций полезности  $U(q;p)$  и спроса и методов оценки их параметров по реально доступным статистическим данным. В более современной литературе, в том числе посвященной моделированию потребительского поведения, идея эндогенных, т.е. выбираемых в процессе покупки, цен не находит отражения [12, 13, 18]. Фактически цена может служить лишь сигналом качества товара в процессе покупки. Но такой подход, традиционно называемый гедоническим, развит и используется в основном для товаров длительного пользования, потребительские свойства которых могут быть хорошо измерены и описаны [12].

В этих условиях представляется важным проанализировать возможности выявления эндогенности цен, по которым домашние хозяйства приобретают товары и услуги, сгруппированные в принимаемые статистическими ведомствами группы. Эндогенность может быть следствием различий в детальных товарных структурах таких групп, реально доступных разным потребителям, дифференциации функций полезности для разных групп потребителей и различной оценки потребителями соотношения «цена-качество» блага. Если гипотеза эндогенности средних цен покупок не будет отвергнута на стадии анализа доступных статистических данных, то задачу разработки теории поведения потребителей, учитывающих это явление, можно будет рассматривать как потенциально актуальную.

### **3. Тестирование гипотезы эндогенности цен на основе доступной статистической информации**

Возможность анализа и моделирования структуры расходов населения критически зависит от доступной статистической информации. При прогнозировании такой

структуры не представляется возможным ограничиться лишь макроэкономическими данными – необходим более точный инструмент. Анализ динамики расходов, как на отдельные продукты, так и на крупные их группы, показывает достаточно сильную волатильность структуры расходов. Для ее учета необходимо привлечение дополнительных микроэкономических характеристик и параметров исследуемого объекта, которые можно получить, используя результаты выборочных обследований.

Для моделирования структуры расходов населения в России на данном этапе используются два основных источника информации – Выборочное обследование бюджетов домашних хозяйств (ВОБДХ), проводимое Росстатом, и Российский мониторинг экономического положения и здоровья населения (RLMS). Эти два источника существенно различаются между собой по способу сбора данных и предоставляемой информации. ВОБДХ является наиболее масштабным обследованием (порядка 50 тыс. наблюдений), а также чаще проводится. Кроме того, его методика сбора информации представляется более обоснованной, хотя и более трудоемкой. В то же время отсутствие полноценного доступа к первичным микроданным ставит ряд серьезных вопросов при попытках использования информации ВОБДХ при построении прикладной модели, в том числе для прогнозирования структуры расходов.

Ключевым с точки зрения моделирования такой структуры является отсутствие в ВОБДХ доступных данных об индивидуальных ценах покупок, необходимых для проверки гипотезы эндогенности цен и для моделирования, в случае принятия этой гипотезы. В то время как такая информация, возможно, собирается в рамках обследования. В данной статье предлагается механизм дополнения информации ВОБДХ, в том числе в части цен покупок, с помощью данных RLMS, в которых такие цены доступны. Для решения вопроса о возможности такого дополнения и способа распространения информации одного обследования на данные другого необходимо сопоставить их структуру, методику проведения, а также параметры самих выборок.

Выборочное обследование бюджетов домашних хозяйств представляет собой комплексное обследование бюджетов, проводимое ежеквартально. В его рамках собираются данные о расходах и доходах семьи<sup>1</sup>, занятости, бытовых условиях, доступе к услугам и товарам, наличии товаров длительного пользования и других показателях. Наиболее важными в контексте поставленной задачи являются данные о покупках и других расходах домашних хозяйств. Такие данные собираются с помощью записей в дневниках и журналах, которые должны вестись на протяжении периода времени (от двух недель до трех месяцев), куда главе домохозяйства следует записывать все расходы, сделанные членами его домохозяйства за отчетный период. Очевидно, что такой способ сбора информации, хотя и не лишен целого ряда недостатков, в том числе допускает возможность недобросовестного заполнения дневников, все же является значительно более полным и информативным, чем, например, разовые опросы о расходах, применяемые в RLMS.

В течение долгого времени Росстат предоставлял доступ только к агрегированным результатам ВОБДХ, которые публиковались в ряде источников. Но затем был открыт свободный доступ всех заинтересованных исследователей к микроданным обследования – информации низкого уровня агрегирования по отдельным домохозяйствам. Такая информация доступна по двум волнам обследования – 2003–2004 гг.

---

<sup>1</sup> В работе отдельные респонденты выборочных обследований считаются семьями – различия между семьями и домохозяйствами представляются несущественными в данном контексте.

(8 кварталов), и 2005–2007 (12 кварталов), которые не вполне сопоставимы между собой из-за различного способа кодирования регионов. У исследователей есть доступ к широкому перечню показателей, для отдельных домашних хозяйств с незначительными колебаниями размера выборки в зависимости от категории данных (46–53 тыс. семей).

Предоставляемые данные являются агрегированными. Так, расходы для отдельных домашних хозяйств собраны в группы продуктов. Например, продовольственные товары представлены лишь по 12 группам, отдельно есть информация о расходах на алкоголь, табачные изделия, некоторые виды непродовольственных товаров (одежда и обувь), некоторые виды услуг (коммунальные услуги), а также несколько видов расходов согласно классификатору расходов по целям (на образование, здравоохранение, досуг и т.п.). С точки зрения моделирования и прогнозирования расходов домашних хозяйств с учетом доходной дифференциации критичным является отсутствие данных об индивидуальных ценах покупок [7].

С точки зрения анализа существенной проблемой также может стать ремонт панели, проводимый Росстатом. Критичным является не столько естественное истощение панели с последующим добавлением новых наблюдений, а присвоение новым домохозяйствам номеров выбывших наблюдений, без явного указания на такую замену. В результате сопоставления данных в разных периодах обнаруживается, что домашнее хозяйство с одним и тем же номером может обладать разными характеристиками, которые из общих соображений должны оставаться относительно стабильными. Так, в любых двух соседних периодах существует достаточно большое число семей (до нескольких процентов), для которых очень значительно отличается размер семьи (на 3–5 человек), а также номер 10-процентной доходной группы, к которой они принадлежат. Это, по-видимому, свидетельствует о замене наблюдения под этим номером на новое.

Возникают существенные проблемы с распространением данных ВОБДХ на всю совокупность семей. Хотя Росстат и предлагает механизм перевзвешивания результатов обследования для более точного их соответствия параметрам генеральной совокупности, даже взвешенные результаты сильно отклоняются от данных макростатистики. Один из базовых параметров – средний доход – отличается на десятки процентов. Более того, даже показатели, рассчитываемые на основе выборочного обследования, впоследствии существенно корректируются. Так, например, коэффициенты дифференциации населения по доходам (коэффициенты Джини и фондов) в агрегированных итогах, приводимых Росстатом [8], достаточно сильно отличается от величины этого показателя, публикуемого Росстатом в официальных справочниках [10].

RLMS, имеющий в своей структуре раздел, посвященный расходам домашних хозяйств на различные товары, является вторым источником информации, который отчасти восполняет пробелы официальной статистики. Обследование RLMS проводится в форме ежегодных интервью представителей домашних хозяйств и, в отличие от ВОБДХ, не подразумевает заполнения дневников или журналов в течение периода времени. Структура вопросов RLMS состоит из нескольких основных разделов, среди которых представлены, в том числе, данные о расходах (раздел E) и доходах (раздел F) домохозяйства. В ряде исследований для выборочных обследований домашних хозяйств было показано, что данные об общих расходах и доходах домашних хозяйств могут довольно значительно расходиться между собой [9]. Так как для анализа требовались в первую очередь данные о расходах на отдельные виды товаров,

то в качестве информации о благосостоянии нами использовалась сумма расходов на все статьи из опросного листа RLMS.

Информация о расходах RLMS включает данные о приобретении продовольственных товаров, расходах на них и количества приобретенных товаров (57 позиций) в натуральном выражении (по 56 позициям, данные о жевательной резинке не имеют натурального выражения). Представлены данные о расходах на целый ряд непродовольственных товаров (одежда: детская и взрослая, техника, топливо и т.д.) и услуг. В силу значительно большей дифференциации непродовольственных товаров, а также более высокой степени агрегированности для них, так же как и для услуг (по естественной причине сложности измерения соответствующего показателя), приводятся данные о покупках лишь в стоимостном выражении.

Для продуктов питания (и только для них) можно получить индивидуальные данные о ценах покупок, что позволяет исследовать зависимость уровня цен покупок от других характеристик, в том числе от уровня благосостояния (дохода).

Данные о расходах на покупку отдельных видов товаров и услуг собираются на основе различных временных горизонтов. Так, расходы на покупку продовольственных товаров фиксируются лишь за период недели, предшествующей интервью. Это вносит значимый случайный фактор – далеко не все (в том числе регулярно потребляемые) продовольственные товары покупаются семьей каждую неделю. Таким образом, целый ряд позиций может быть опущен в обследовании при заполнении домашнего хозяйства соответствующего журнала в силу малой длины временного интервала. Такие пропуски приводят к значительным различиям структур покупок для каждой из семьи в соседних волнах опроса и значительно затрудняют межвременные сопоставления.

Покупки большинства непродовольственных товаров фиксируются на протяжении периода в три месяца; для услуг этот горизонт, как правило, равен одному месяцу. При такой форме опроса (ежегодное интервью) автоматически возникают два эффекта, влияющие на точность получаемых данных о расходах. Так, увеличение рассмотренных выше горизонтов ведет к неполноте данных в силу того, что интервьюируемый может просто не точно помнить все свои расходы несколько месяцев назад и упустить некоторые существенные траты. Недостаточная длина такого горизонта ведет к пропуску некоторых позиций просто в силу нерегулярности покупок каждой из них, кроме того, может иметь значение фактор сезонности.

Следует иметь в виду также отсутствие вопросов об общей сумме расходов на те или иные категории товаров и раздела «Прочие расходы», как в общем, так и по категориям (например, прочие расходы на продовольственные товары). Хотя список вопросов представляется вполне исчерпывающим, наличие таких позиций также могло бы нести некоторую дополнительную информацию. Категория «питание вне дома» представлена очень кратко. Отдельно фиксируется информация о расходах на эти услуги за неделю. Но суммарные данные о расходах на питание дома и вне дома (уже за месяц) сильно расходятся с суммой соответствующих статей из вопросника (расхождение может быть в обе стороны, без системы), что говорит о явной неточности собранных данных.

В процессе анализа данных RLMS были выявлены две основные проблемы, которые могут повлиять на результаты и точность оценивания параметров любых моделей. Первая – неполная система ответов. То есть ситуация, когда опрошиваемый указывал, что приобретал тот или иной продукт питания, но «забывал» указать количество или стоимость (или, реже, и то, и другое). Такое искажение занижает общую

сумму расходов на не очень значительную величину (порядка одного или нескольких процентов). Вторая проблема, в отличие от первой, носит более содержательный характер (тогда как первая – скорее технический) и может быть названа «рыхлостью данных». Многие домашние хозяйства, особенно хозяйства с относительно низким уровнем доходов/расходов, за отчетный период приобретали лишь небольшую часть номенклатуры всех товарных позиций в опросе – в среднем 15 позиций (27% от общего числа позиций номенклатуры). Такое поведение, возможно, связано с коротким горизонтом опроса – одна неделя. «Рыхлость» и наличие явных неточностей в индивидуальных данных RLMS (в том числе пропуски в ответах) не позволили работать напрямую с исходными данными по отдельным домашним хозяйствам для анализа взаимосвязи уровня доходов/расходов и цен покупок.

Было выдвинуто предположение о зависимости цен покупок от уровня доходов. Для более корректного определения уровня цен покупок в рамках такого предположения и в ситуации описанных проблем с расчетом индивидуальных цен обоснованным выглядит расчет групповых индексов цен. Предлагается использовать цены покупок, единые для некоторых групп домашних хозяйств с близким уровнем расходов/доходов. Выделение таких групп может осуществляться различными способами. В данном случае предлагается простой механизм, который подтвердил свою состоятельность при применении к данным разных лет. Выделяется заранее определенное число равных по количеству домохозяйств групп на основе ранжирования наблюдений в выборке по уровню целевого показателя. Для целей получения индексов цен покупок продуктов питания таким показателем был выбран уровень расходов на питание в домохозяйстве в пересчете на одного человека.

Предполагается, что выделяемые таким образом группы потребителей являются относительно однородными и представительными для рассматриваемой задачи выявления дифференциации цен покупок по укрупненным позициям продовольственных товаров в зависимости от показателей, определяющих их потребительское поведение. В числе достоинств данного механизма стоит назвать простоту реализации и гибкость, позволяющие легко менять число групп, выбирать показатель, по которому происходит ранжирование, а также применять его к выборкам любого размера. В процессе моделирования рассматривались и другие показатели, по которым проводилось деление на группы – в том числе общая сумма расходов. Также проводилось варьирование границ групп для исследования вопроса чувствительности выделенных таким способом групп к небольшим отклонениям в их структуре. Общим выводом является достаточно высокая устойчивость результатов к подобным изменениям в процедуре выделения групп.

В проведенном анализе по предлагаемой методике использовались данные 16-й волны RLMS за 2007 г. Наиболее новыми данными RLMS на момент написания статьи были результаты опроса за 2008 г., но для корректности последующего сопоставления с данными ВОБДХ, доступными лишь по 2007 г., моделирование велось для данных RLMS также за 2007 г. Для проверки корректности результатов проведенные вычисления были повторены для соседних периодов (2006, 2008 гг.), и результаты оказались стабильными – отмеченные тенденции и закономерности подтверждаются в данных за соседние периоды.

После исключения незначительного количества наблюдений, не предоставивших данных о своих расходах за отчетный период, исследуемая совокупность составила 5410 домохозяйств с ненулевой суммой расходов на питание. Для исследования стабильности полученных результатов указанным способом – через ранжирование по



сумме расходов на питание на одного члена домохозяйства – было сделано несколько разбиений на группы данных RLMS за 2007 г.:

- малое число групп – 10 (по 540 домашних хозяйств в группе);
- среднее число групп – 20 (по 270 домашних хозяйств в группе);
- большое число групп – 50 (по 106 домашних хозяйств в группе);
- наибольшее число групп – 100 (по 53 домашних хозяйства в группе).

Во всех вариантах разбиения по каждой группе были рассчитаны показатели суммарных расходов на покупку каждого вида продуктов и их суммарное количество. Средние цены были рассчитаны как отношение общей суммы расходов на этот товар в группе к приобретенному количеству.

На основе этих показателей были рассчитаны четыре основных индекса цен: Ласпейреса (IPL), Пааше (IPP), Фишера (IPF) и Монтгомери (IPM). Эти индексы могут интерпретироваться как «пространственные», так как все относятся к одному периоду времени, но сравнивают между собой уровень цен в разных группах домашних хозяйств.

Рассмотрим последовательно четыре случая разбиения на группы.

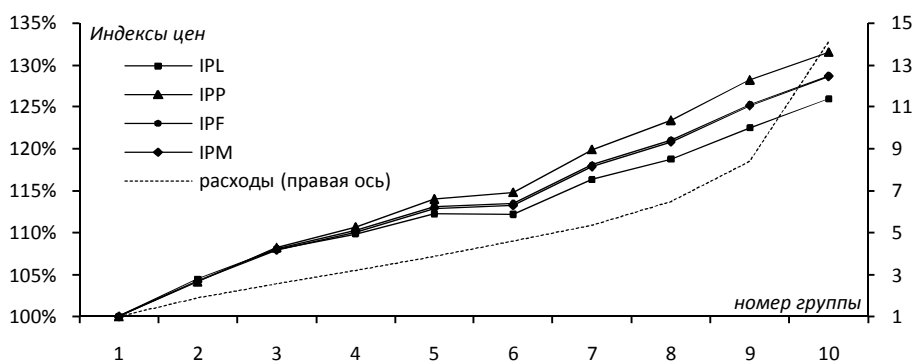


Рис. 1. Зависимость цен покупок от суммы расходов на питание по 10 группам

На рис. 1 изображена динамика четырех индексов в базовом исчислении для случая разбиения на 10 групп. Под динамикой понимается не отношение цен в разные периоды времени (так как все моделирование ведется для одного периода), а изменение средних цен покупок продуктов питания при движении от групп с меньшими доходами к группам с большими. Базисные индексы рассчитаны по отношению к первой группе (с самыми низкими расходами на питание) путем последовательного перемножения цепных индексов. Относительное значение расходов отложено по правой оси.

Отметим несколько важных фактов. Все четыре индекса возрастают, т.е. при малом количестве групп гипотеза о том, что больший доход (в данном случае совокупные расходы на одного члена домохозяйства) ведет к более высоким средним ценам покупок, в целом подтверждается. Уровень расходов на продукты питания растет значительно большими темпами, чем межгрупповые индексы цен. Коэффициент фондов (отношение суммы подушевых расходов на питание 10% самых «богатых» к 10% самых «бедных») составляет 14 раз, а цепные индексы цен на продукты питания

на таком диапазоне различаются в среднем лишь на 29%. Такое различие свидетельствует в пользу того, что рост суммы расходов не может быть объяснен только дифференциацией цен покупок по рассматриваемым группам домашних хозяйств.

Гипотетически возможна такая ситуация, что все домашние хозяйства в выборке потребляют примерно одинаковые количества товаров, но по разным ценам, и все различия в сумме расходов на продукты питания могут объясняться только разными ценами покупок. Тогда положительная взаимосвязь между уровнем расходов и цен была бы автоматической и очень просто свидетельствовала в пользу подтверждения исследуемой гипотезы. Однако, как уже было отмечено, различия в расходах несопоставимо больше, чем различия в ценах покупок, и анализ структуры расходов показывает, что более богатые семьи, действительно, потребляют значительно больше продовольственных товаров.

Межгрупповые индексы цен Фишера (IPF) и Монтгомери (IPM) очень близко, практически неразличимо зависят от душевого расхода на продовольственные товары, что хорошо согласуется с теорией и косвенно подтверждает корректность построения индексов. Нетипичной является ситуация, что IPP имеет завышенную динамику по сравнению с IPL.

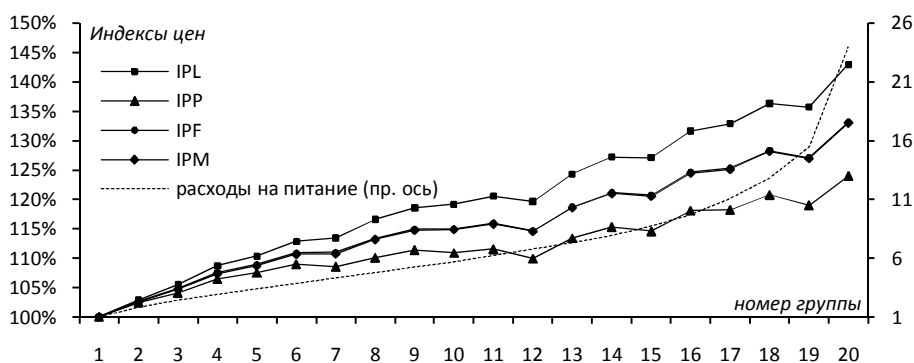


Рис. 2. Зависимость цен покупок от суммы расходов на питание по 20 группам

На рис. 2 изображена аналогичная последовательность межгрупповых базисных индексов для 20 расходных групп. В целом, уровень цен, рассчитанный по каждому индексу, растет с ростом расходов, что также подтверждает исследуемую гипотезу. Однако в данном случае рост уже не является монотонным — имеются некоторые группы, для которых происходит снижение индекса. Это означает, что уровень цен в следующей группе (с большей суммой расходов на питание) ниже, чем в предыдущей (с меньшей суммой расходов), что, видимо, объясняется различием других характеристик домашних хозяйств в соседних группах. Общий рост уровня цен даже несколько выше, чем в случае 10 групп (33 против 29% по индексам цен Фишера и Монтгомери). Также различные индексы дают несколько больший разброс.

Рост общего уровня расходов по-прежнему значительно опережает по темпам рост уровня цен по выделенным группам для рассматриваемого периода. Более того, он значительно ускоряется с увеличением числа групп, на которые разбивается интервал. Пятипроцентный коэффициент фондов составляет уже 24.

Индексы цен Монтгомери (IPM) и Фишера (IPF), как и в предыдущем случае, имеют очень близкую, практически неразличимую динамику. В данном случае индекс Ласпейреса (IPL) значительно опережает индекс Пааше (IPP), что согласуется с описанными выше наблюдаемыми особенностями в данных.

В такой ситуации корректный выбор индекса цен является важным этапом последующего анализа. Хотя общие качественные результаты являются довольно общими, межгрупповая динамика чувствительна к выбранному индексу цен и результаты количественного моделирования могут сильно зависеть от этого выбора.

В результате сопоставления динамики базисных «пространственных» индексов цен покупок продуктов питания и динамики общей суммы расходов на питание для различного числа групп была установлена достаточно устойчивая зависимость между этими показателями<sup>2</sup>. Кроме общего вывода о наличии указанной зависимости было выявлено несколько характерных тенденций, которые проявились при переходе от разбиения на меньшее число групп к большему. С ростом числа групп:

- незначительно ускоряется средний рост уровня цен конечного периода по отношению к начальному (по индексам IPF и IPM);
- увеличивается «изрезанность» динамики цен – растет частота случаев уменьшения уровня цен при росте дохода, которая может быть связана с наличием случайного фактора при отнесении домохозяйства к расходной группе в силу особенностей сбора данных;
- растет расхождение между уровнями цен, рассчитанными по разным индексам (в том числе индексам IPM и IPF);
- замедляется (вплоть до отрицательной) динамика уровня цен по индексу цен Пааше, т.е. при большом числе групп и высокой волатильности наблюдается тенденция к уменьшению уровня цен покупок (по IPP) с ростом расходов;
- значительно ускоряется «динамика» цен по межгрупповому индексу цен Ласпейреса (IPL).

Для подтверждения выявленных зависимостей и уточнения результатов были построены несколько регрессионных моделей, которые во многом подтверждают полученные выводы.

Исследовалась зависимость уровня цен покупок от общей суммы душевых расходов. В этой модели уравнения принимали достаточно простой вид – зависимость индекса цен от отношения логарифмов расходов в каждой группе к первой<sup>3</sup>:

$$(1) \quad IPF = -0,3 + 1,185 \cdot Expenditures, R^2 = 0,91, \\ (-4,36) \quad (21,84)$$

$$(2) \quad IPM = -0,46 + 1,347 \cdot Expenditures, R^2 = 0,92, \\ (-6,35) \quad (23,79)$$

где  $Expenditures = \ln(\text{душевые расходы на питание } k\text{-й группы}) / \ln(\text{душевые расходы на питание } 1\text{-й группы})$ .

<sup>2</sup> Расчеты для случаев разбиения на 50 и 100 групп не приводятся в данной статье, но могут быть предоставлены автором.

<sup>3</sup> Под коэффициентами регрессионных уравнений указаны значения соответствующих t-статистик.

Подобные уравнения были построены для индексов Фишера и Монтгомери для случаев разбиения на 50 (уравнения (1) и (2)) и 100 групп (не приводятся). Отметим, что все четыре регрессии (оба индекса цен для случая 50 и 100 групп) характеризуются высокими значениями показателя  $R^2$  (90–94%), значимыми оценками коэффициентов  $\beta$  (при переменной *Expenditures*), несколько большими единицы.

Кроме того, были рассчитаны вспомогательные регрессии индекса цен Фишера на индекс цен Монтгомери для 50 групп:

$$(3) \quad IPF = -0,1 + 1,13 \cdot IPM, R^2 = 0,998. \\ (-11,8) (154,1)$$

Моделирование для случая 50 и 100 групп дает очень близкие результаты, что подтверждает устойчивость исследуемой зависимости. Для обоих рассмотренных индексов цен были получены уравнения со значимыми положительными коэффициентами.

Значение коэффициента  $\beta$  в уравнении при использовании индекса Фишера в обоих случаях превышает значение этого коэффициента при использовании индекса Монтгомери, т.е. IPF характеризуется несколько более сильной (по сравнению с регрессией для IPM) зависимостью от соотношения логарифмов душевых расходов в выбранной группе и группе с наименьшими такими расходами. Это также подтверждается результатами вспомогательных регрессий IPF на IPM, в которых коэффициент перед IPM также больше единицы. Вспомогательные регрессии, характеризующиеся очень высокими  $R^2$  (98–99%), подтверждают отмеченную тесную связь индексов цен Фишера и Монтгомери применительно к данному исследованию. В то же время с увеличением числа групп коэффициент наклона в регрессии IPF на IPM вырос с 1,13 до 1,71, что связано с увеличением волатильности индексов Ласпейреса и Пааше, лежащих в основе расчет IPF.

Устойчивая положительная зависимость между ценами покупок и уровнем расходов (доходов) может отчасти объясняться взаимным влиянием цен продуктов на уровень расходов на питание и обратно. В домохозяйствах с высокими подушевыми расходами детальная структура покупок продуктов питания может быть смещена в сторону более качественных и дорогих товаров. Для оценки степени такого обратного влияния был проведен аналогичный анализ, в котором наблюдения ранжировались не по сумме расходов на питание, а по общей сумме расходов на одного члена домохозяйства. В этот показатель включаются все расходы по анкете, в том числе расходы на продовольственные товары, питание вне дома, а также покупки по всему перечню непродовольственных товаров, услуг (в том числе коммунальных) и т.д., за исключением сбережений (в любой форме). Возможно, в данной ситуации было бы лучше использовать показатель доходов. Однако в научной и методической литературе не раз высказывалась точка зрения, согласно которой достоверность такого показателя в опросах невысока.

Показатель общих расходов также зависит от уровня цен, так как расходы на продукты питания входят в общую сумму. Представляется, что это влияние невелико. Так, средняя доля расходов на продукты питания в общей сумме расходов составляет 37%.

Отметим, что с переходом к общей сумме расходов на человека основные наблюдаемые эффекты и, как следствие, выводы остаются без изменений. В целом, прослеживается тенденция к увеличению уровня цен в группах с большей суммой

общих расходов (доходов). Во многих случаях рост цен с увеличением дохода происходит даже быстрее (для 10, 20 и 50 групп) и несколько более плавно (например, для 20 групп).

Значительно отличается межгрупповая динамика расходов на продукты питания. Но так как ранжирование проводилось не по этому показателю, то ее поведение не столь гладкое и увеличение суммы расходов в старших доходных группах не столь быстрое. В такой ситуации использование ранжирования и выделения групп на основе душевых расходов на питание представляется вполне оправданным и более корректным при моделировании цен покупок продуктов питания. Цены на товары этой категории, по-видимому, должны быть наиболее тесно связаны именно с суммой расходов на питание, что позволяет остановиться на этом показателе для проверки выдвинутой гипотезы.

При увеличении числа групп (дроблении выборки) результаты становятся менее устойчивыми – растет волатильность индексов цен и расхождение между разными индексами. Данное обстоятельство еще раз подтверждает, что работа с отдельными домашними наблюдениями в выборке не является корректной – необходимо агрегирование семей в типичные группы. Но основные выводы сохраняются – базисные индексы Фишера и Монтгомери цен покупок значимо растут с увеличением расходов на питание.

Анализ показал, что существует зависимость между средними ценами покупок продуктов питания и благосостоянием семьи. В рамках сделанных предположений уровень дохода характеризовался суммой душевых расходов на покупку продуктов питания или суммой общих расходов на члена домохозяйства. При разбиении на расходные группы средние цены покупок продуктов питания растут по мере увеличения общей суммы расходов (доходов). То, что цены покупок связаны именно с доходной характеристикой домохозяйства (в данном случае – суммой расходов на питание или общей суммой расходов), свидетельствует в пользу того, что потребители выбирают цены, по которым приобретают те или иные товары, учитывая возможные некоторые скрытые параметры (качество, сервис и т.д.)

В контексте выдвинутой гипотезы эндогенности цен, т.е. выбора цен покупок при принятии потребительских решений, обнаруженная зависимость между средними ценами покупок и уровнем благосостояния – решающий аргумент в пользу принятия этой гипотезы. Анализ доступных микроэкономических данных показывает, что цены покупок не являются одинаковыми и заданными внешне для разных категорий населения.

#### 4. Соединение информации из двух выборок

В условиях обнаруженной взаимосвязи цен покупок и микроэкономических характеристик домашних хозяйств ключевым является учет этих эффектов при моделировании и прогнозировании структуры спроса населения на отдельные продукты и группы товаров. Для такого моделирования могут быть использованы обследования RLMS и ВОБДХ. Проводимое Росстатом обследование является более точным инструментом в силу большего объема выборки, более высокой частоты проведения и ряда других факторов. Ключевым его недостатком, в контексте принятия гипотезы эндогенности цен, является отсутствие информации об индивидуальных ценах покупок даже для отдельных категорий товаров (например, продуктов питания). В резуль-

тате построение модели, учитывающей дифференциацию цен, на базе только этих данных становится невозможным.

В этой связи актуальной является разработка механизма, позволяющего совместно учитывать информацию из двух разных обследований для получения всех необходимых для моделирования параметров предполагаемых зависимостей. Так как ВОБДХ больше подходит для построения детальной модели и получения прогноза, необходимо проведение процедуры, которая позволит дополнить информацию этого обследования имеющимися данными о ценах покупок продуктов питания в выборке RLMS. В таких условиях хотя бы часть эффектов, связанная с дифференциацией потребительского поведения в части выбора цен покупок продуктов питания, будет учтена при получении прогноза структуры спроса.

Для решения вопроса о возможности распространения информации о ценах покупок по доходным группам из результатов RLMS на данные ВОБДХ сравним эти выборки по основным параметрам. Для такого сравнения были выбраны следующие показатели, которые могут быть получены из обеих выборок:

- средняя по каждой группе сумма подушевых расходов на питание;
- средняя по каждой группе сумма общих потребительских расходов на человека;
- средний размер семьи в каждой группе;
- средняя доля расходов на питание в общей сумме расходов в каждой группе;
- доля городского населения в группах.

Опишем эти показатели. Один из крупных агрегатов, имеющихся в базовом наборе индивидуальных данных ВОБДХ, – сумма расходов семьи на питание за отчетный период, не включающая расходы на алкоголь, которые выделены отдельно, и табачные изделия, учитываемые в непродовольственных расходах. Этот показатель был рассчитан на одного члена домохозяйства. Для получения аналогичного показателя по данным RLMS были сложены все подушевые расходы на продукты питания, за исключением алкогольных напитков и табачных изделий, которые в этом обследовании отнесены к продуктам питания.

Другим важным показателем ВОБДХ является общая сумма потребительских расходов, состоящая из покупок продуктов питания, алкоголя, непродовольственных товаров, затрат на услуги и питание вне дома. Для сопоставления с ним по данным RLMS был рассчитан аналогичный показатель, включающий в себя все расходы на покупку товаров (в том числе продукты питания, непродовольственные товары, топливо и т.д.) и оплату услуг (коммунальных, бытовых, образования, здравоохранения и др.), но без учета сбережений, отчислений на обязательные платежи, помощь родственникам и т.п. Обе эти суммы расходов также были приведены к расходам на одного члена домохозяйства.

Доля расходов на питание в общей сумме расходов представляет собой среднее значение по группе из отношений первых двух показателей для каждого домохозяйства.

Средний размер семьи – это данные о количестве человек в семье в каждой из выборок без разделения на взрослых и детей, так как такое разделение достаточно трудно выполнить по данным RLMS.

В качестве доли городского населения в группах домохозяйств из ВОБДХ использована информация о проживании семьи в городском или сельском населенном пункте, которая в явном виде представлена в данных обследования. В RLMS пред-

лагается несколько более детальное деление населенных пунктов по типам. Для сопоставления с результатами ВОБДХ областные центры и города были отнесены к городскому населению, а поселки и села – к сельскому. Такое деление представляется оправданным как содержательно, так и статистически близко к значениям ВОБДХ, что будет показано ниже.

Дальнейший анализ проводился как для средних по году значений, так и отдельно по IV кварталу, который по своим параметрам наиболее близок к параметрам RLMS за 2007 г. В качестве годовых данных для домохозяйств рассматривались суммарные показатели расходов за четыре квартала, если речь идет о показателях типа поток, или средние, если рассчитывались параметры типа запас, например размер семьи или доля городского населения.

На следующем этапе были рассчитаны средние по группам значения этих пяти ключевых показателей (для 5-процентных и 10-процентных групп), а также выборочные значения стандартных отклонений каждого из параметров по группам, отдельно для RLMS и ВОБДХ за 2007 г. Сопоставление показывает, что распределения всех этих показателей по группам очень близки для двух выборок (рис. 3). На графиках представлена межгрупповая динамика средних по группам значений пяти рассматриваемых параметров двух выборок, по горизонтальной оси – номер группы по расходам на питание на человека.

Во всех случаях наблюдается повторение качественных тенденций в распределении каждого из показателей с ростом номера расходной группы, в том числе немонотонной «динамике», причем как по средним значениям, так и по рассчитанным значениям стандартных отклонений.

Денежные расходы в ВОБДХ без использования взвешивания, предлагаемого Росстатом, как правило, несколько ниже, чем аналогичные показатели в RLMS за тот же период. Прокомментируем обнаруженные закономерности в межгрупповой динамике рассматриваемых параметров с ростом расходной группы. Они, по-видимому, достаточно характерны в силу того, что проявляются в обеих выборках. Полученная динамика во многих случаях выглядит весьма стабильной. Для 10-процентных групп почти во всех случаях тенденции являются монотонными, изредка наблюдается плавная смена тенденций, резких скачков или колебаний не отмечается. Для 20 групп поведение показателей несколько менее регулярное, но также не отмечается сильных колебаний или значительных отклонений от тренда (рис. 3). Для ВОБДХ динамика значительно более гладкая, чем для RLMS, что вполне объяснимо большим числом наблюдений.

С ростом душевых расходов на питание в семье естественно увеличивается и общая сумма расходов, а также, постепенно замедляясь, растет доля городского населения в старших расходных группах.

Отметим интересный эффект. Доля расходов на питание растет с увеличением суммы душевых расходов на эту статью. Общая сумма расходов, как было отмечено, также положительно коррелирована с уровнем расходов на питание. В то же время для обеих выборок справедливо оказывается выполнение так называемого закона Энгеля, т.е. уменьшение доли расходов на питание в общей сумме потребительских расходов. Таким образом, между этими тремя показателями возникает нетривиальная связь, когда два из них положительно коррелируют между собой, а третий – положительно с одним из них, но отрицательно с другим. Такая ситуация может объясняться двойственной природой доли расходов на питание. С одной стороны, она увеличивает

ся с ростом числителя, т.е. суммы расходов на питание, с другой – падает с ростом знаменателя – общей суммы расходов.

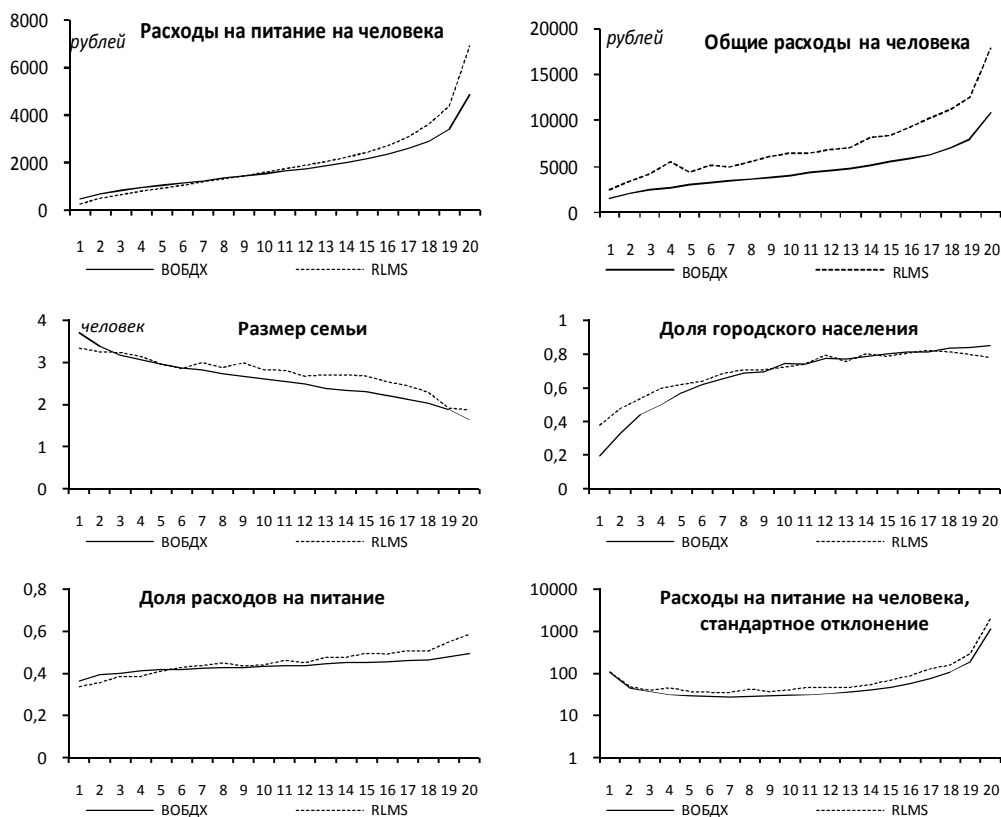


Рис. 3. Сравнение основных параметров выборок RLMS и ВОБДХ для 20 расходных групп

Межгрупповая динамика выборочных стандартных отклонений в зависимости от доходной группы выглядит несколько менее регулярной, но при этом также на качественном уровне, а часто и численно, очень близка в обеих выборках. Отметим незначительное сокращение разброса числа человек в домохозяйстве с ростом номера расходной группы и резкое увеличение разброса уровня душевых расходов на питание на левом и правом концах распределения. Даже в логарифмическом масштабе видно очень существенное увеличение значения выборочного стандартного отклонения внутри группы для младших и, особенно, старших групп (рис. 3).

В рамках поставленной задачи необходимо на основе доступной информации о параметрах обеих выборок разработать и реализовать механизм, позволяющий с помощью оценок «восстановить» динамику цен на продукты питания для данных ВОБДХ. Такой механизм позволит использовать данные RLMS о ценах покупок совместно с данными ВОБДХ, проведя соответствующую корректировку на различие в параметрах выборок. Структура выборок представляется очень близкой друг к



другу, что дает основания для проведения такого оценивания, в то же время корректировка на различия параметров может сделать оценку индексов цен покупок еще более надежной.

Такая «пересадка» индексов цен покупок продуктов питания может осуществляться несколькими принципиально разными путями. Простой заключается в копировании и использовании этих индексов, рассчитанных на основе данных RLMS в выборке ВОБДХ. Такое заимствование представляется оправданным в силу близости характеристик выборок, но лишь на качественном уровне. То есть можно утверждать, что для ВОБДХ также будет характерна тенденция увеличения средних цен покупок продуктов питания с ростом суммы расходов на питание или общей суммы расходов. Сохранятся и другие выводы, которые сделаны для данных RLMS. Но такое простое заимствование индексов цен не учитывает существующие, пусть и незначительные, различия в параметрах выборок, проявляющиеся в несколько разной межгрупповой динамике исследуемых показателей.

Рассмотрим более продвинутый способ использования заимствованных индексов из RLMS для выборки ВОБДХ. Для этого необходим механизм корректировки индексов цен на пусть и не значительные, но существующие отличия в характеристиках расходных групп в двух выборках. Для реализации такого механизма была оценена регрессионная модель индекса цен покупок продуктов питания по группам населения, выделенным по уровню расходов на питание, в зависимости от всех доступных параметров выборки. В данном случае необходимо учитывать характеристики, которые доступны для обеих групп. Используются те пять характеристик, по которым проводилось сравнение выборок по структуре: средняя по группе сумма расходов на питание на человека; средняя по группе сумма общих потребительских расходов на человека; средний размер семьи; средняя доля расходов на питание в общей сумме расходов по группе; доля городского населения в группах.

Необходимо разложить динамику индекса цен по пяти рассмотренным факторам. Основной сложностью в этой задаче является крайне ограниченное число наблюдений при работе с достаточно крупным разделением выборки (10–20 расходных групп), не позволяющее использовать многие стандартные подходы и тесты, применимые лишь для случая большого числа групп. Возможен вариант перехода к более дробным группам домохозяйств (например, 50), но в данном случае в межгрупповой динамике показателей, в том числе индексов цен, появляется достаточно сильная волатильность.

Рассмотрим подробнее парные корреляции отдельных показателей между собой и с индексом цен Фишера, демонстрирующим среднюю динамику из рассмотренных четырех индексов, на данных RLMS (табл. 1). Заметим, что все значения коэффициентов корреляций крайне велики по модулю – более 0,65, а многие – более 0,9. Кроме того, знаки корреляций весьма регулярны. Единственный показатель, который отрицательно связан с другими, – размер семьи. Он имеет отрицательные корреляции со всеми четырьмя факторами и с межгрупповой динамикой цен. Все остальные корреляции между факторами и индексом цен – положительны. Такая ситуация может быть, в том числе, связана с наличием довольно устойчивой тенденции в межгрупповой динамике каждого из факторов и индекса цен – отрицательной для размера семьи и положительной для всех остальных. Знаки корреляций важны, в том числе, для построения так называемых «гармонических регрессий» [4].

Таблица 1.

**Значения парных корреляций факторов  
и IPF на данных RLMS для 20 групп**

	FOOD_EXP	FAM_SIZE	URBAN	FOOD_SH	TOTAL_EXP	IPF
FOOD_EXP	1,00	-0,95	0,67	0,92	0,99	0,91
FAM_SIZE	-0,95	1,00	-0,75	-0,95	-0,95	-0,94
URBAN	0,67	-0,75	1,00	0,88	0,71	0,88
FOOD_SH	0,92	-0,95	0,88	1,00	0,93	0,98
TOTAL_EXP	0,99	-0,95	0,71	0,93	1,00	0,93
IPF	0,91	-0,94	0,88	0,98	0,93	1,00

Для каждого из разбиений на группы (мы рассматриваем выделение 20-ти или 50-ти групп домохозяйств по душевым расходам на питание указанным выше способом) возможно оценивание нескольких вариантов моделей. Так, основным различием является функциональная форма – оценивались обычная линейная регрессия и линейная в логарифмах модели. Для пояснения процесса выбора модели и отбора значимых факторов приведем модели общего вида, включающие полный набор факторов:

$$(4) \quad IPF = 0,86 - 1,41 \cdot 10^{-5} \cdot Food\_exp - 0,03 \cdot Fam\_size + \\ (3,92) \quad (0,59) \quad (-0,79) \\ + 0,21 \cdot Urban + 0,33 \cdot Food\_sh - 1,52 \cdot 10^{-5} \cdot Total\_exp, R^2 = 0,975, \\ (1,98) \quad (0,79) \quad (1,76)$$

где *Food\_exp* – средняя по группе сумма подушевых расходов на питание в домохозяйстве; *Fam\_size* – среднее по группе число человек в домохозяйстве (взрослых и детей); *Urban* – доля домохозяйств в группе, проживающих в городах; *Food\_sh* – средняя по группе доля расходов на питание в домохозяйстве; *Total\_exp* – средние по группе общие расходы домохозяйства.

$$(5) \quad \ln(IPF) = - 0,66 + 0,04 \cdot \ln(Food\_exp) + 0,01 \cdot \ln(Fam\_size) - \\ (-1,18) \quad (0,57) \quad (0,09) \\ - 0,02 \cdot \ln(Urban) + 0,12 \cdot \ln(Food\_sh) + 0,06 \cdot \ln(Total\_exp), R^2 = 0,971. \\ (-0,23) \quad (0,53) \quad (1,36)$$

Прокомментируем подробнее полученные регрессии. Видно, что результаты применения моделей (4), (5) очень близки. Значение  $R^2$  в обеих моделях велико (более 97%) и подтверждает гипотезу о том, что выбор цен покупок (индекс относительно первой группы) во многом определяется микроэкономическими характеристиками группы домохозяйств. Для случая 50-ти групп результаты качественно не отличаются, значения  $R^2$  несколько ниже, но все равно находятся на уровне выше 93%.

Межгрупповая динамика индекса цен по данным RLMS, рассчитанная с помощью обеих моделей, приводится на рис. 4. Однако говорить о преимуществах одной из них на этом этапе не вполне корректно.

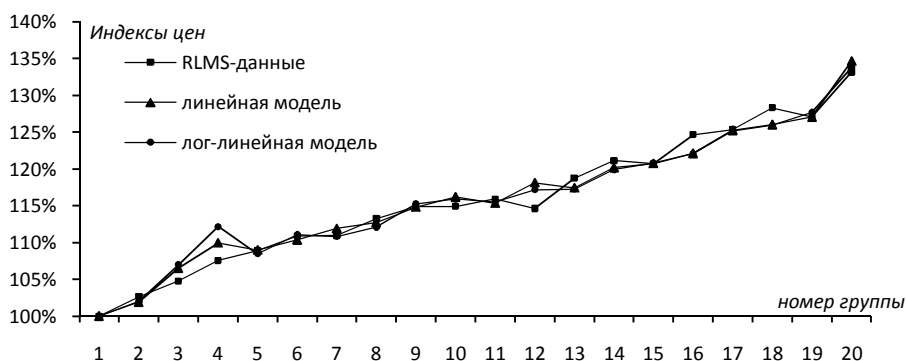


Рис. 4. Результаты регрессионного моделирования межгруппового индекса цен Фишера по данным RLMS для 20 групп

Обратим внимание на незначимость коэффициентов в моделях. При столь малой длине выборки все параметры в случае линейной в логарифмах и почти все в случае линейной моделей оказываются незначимыми. Одной из причин этого может являться сильная мультиколлинеарность факторов. Это подтверждается при оценивании вспомогательных регрессий одного из факторов на другие – во всех случаях значение  $R^2$  оказывается крайне высоким (до 98–99%). В такой ситуации незначимые коэффициенты должны быть исключены из модели, так как включение стольких тесно коррелированных факторов делает модель некорректной и искажает получаемые оценки. Однако простое «механическое» исключение незначимых факторов в порядке наименьшего уровня значимости в данном случае не вполне применимо. Необходимо учитывать содержательные связи между ними для получения наиболее точных оценок.

После исключения незначимых коэффициентов линейная модель включает в себя общие расходы и долю городского населения, а логлинейная – общие расходы и долю расходов на питание. При 50 группах в обеих моделях оказались значимыми по три фактора – общие расходы, доля расходов на питание, расходы на питание. С одной стороны, в такой ситуации не полностью учитывается различие между выборками. Все пять показателей имеют свою межгрупповую динамику для каждой из выборок. Идея пересчета индекса цен предполагает учет максимально возможной информации при восстановлении индексов для ВОБДХ по параметрам выборки. Но, с другой стороны, имея в виду очень тесную корреляцию между факторами, есть возможность подобрать такое их сочетание, которое будет объяснять значительную долю колебаний индекса цен и в то же время не будет характеризоваться столь значительной мультиколлинеарностью.

Одним из возможных способов решения этой проблемы может стать применение метода главных компонент (МГК) для линейной модели, учитывающих влияние всех факторов. Были рассчитаны такие линейные комбинации факторов, составля-

ющие первые две главных компоненты для случая 50-ти групп и одну для случая 20-ти групп, и оценены регрессии индекса цен Фишера на них. Результаты моделирования являются приемлемыми, хотя значение  $R^2$  оказалось несколько ниже, чем в обычных моделях – порядка 82–86%. Такая модель дает несколько заниженную межгрупповую динамику индекса цен по сравнению с обычными регрессиями и чувствительна к выбору количества главных компонент, включаемых в уравнение. В ситуации затрудненной интерпретации полученных таким способом результатов было решено не использовать этот метод в качестве основного.

Окончательный выбор модели должен быть основан не только на априорном выборе факторов, но и на результатах применения модели к расчету межгрупповой динамики цен для ВОБДХ.

Вторым этапом предлагаемого механизма оценки межгрупповых индексов цен для данных ВОБДХ по информации RLMS становится восстановление таких индексов с помощью полученных регрессионных моделей и параметров выборки ВОБДХ. Одной из основных сложностей на этом этапе стало большое разнообразие полученных моделей в силу разного вида функциональной формы, использования метода главных компонент, а также многочисленных возможных корректировок, описанных ниже. Общее количество вариантов моделей в этом случае оказалось порядка двух десятков для каждого числа групп. В такой ситуации затруднительно выбрать одну наиболее точную и универсальную модель, которую можно предложить в качестве искомого механизма «пересадки» индексов.

Итак, чтобы оценить межгрупповые индексы цен для ВОБДХ, в рассчитанные регрессионные уравнения для соответствующего числа групп должны быть подставлены параметры ВОБДХ каждой группы. Но простое восстановление индексов цен таким способом может оказаться не вполне корректным. Например, рассчитанный индекс цен для первой группы (рассматриваются базисные индексы, отнесенные к младшей группе) может не быть равным единице, и возможно ввести соответствующую нормировку, разделив все полученные индексы на его значение для первой группы.

Другой возможной корректировкой может стать включение в модель регрессионных остатков, рассчитанных в момент разложения индекса цен по пяти факторам по данным RLMS. Преимущество такой корректировки заключается в автоматическом учете направления и силы возникающего смещения, которое повторяется в рассматриваемой выборке ВОБДХ, как и в исходной RLMS.

Приведем для примера график (рис. 5) межгрупповой динамики исходных данных RLMS и шести различных рассчитанных значений для ВОБДХ с помощью линейной модели (где «норм» – с учетом нормировки на значение индекса в первой группе, «остатки» – включение остатков в модель, «МГК» – модель, рассчитанная с помощью метода главных компонент для первой компоненты) для случая разбиения на 20 групп.

Видно, что межгрупповая динамика достаточно сильно отличается для разных вариантов модели, даже в рамках линейной функциональной формы. Так, например, не использование нормировки на значение индекса в первой группе приводит к достаточно сильному отклонению расчетного значения от единицы, что затрудняет интерпретацию. Если для обычной регрессии это приводит к занижению значения индекса в первой группе, то для регрессии с помощью метода главных компонент – к завышению. Обычные регрессии, построенные без применения МГК, нормированные на значения индекса в первой группе, сильно «завышают» межгрупповую динамику

индексов цен, в том смысле, что для всех групп значение базисного индекса оказывается выше, чем для исходных данных RLMS. Использование предлагаемой корректировки регрессионными остатками позволяет точнее повторить отдельные колебания в межгрупповой динамике, например, некоторый провал, наблюдаемый в 12-й группе.

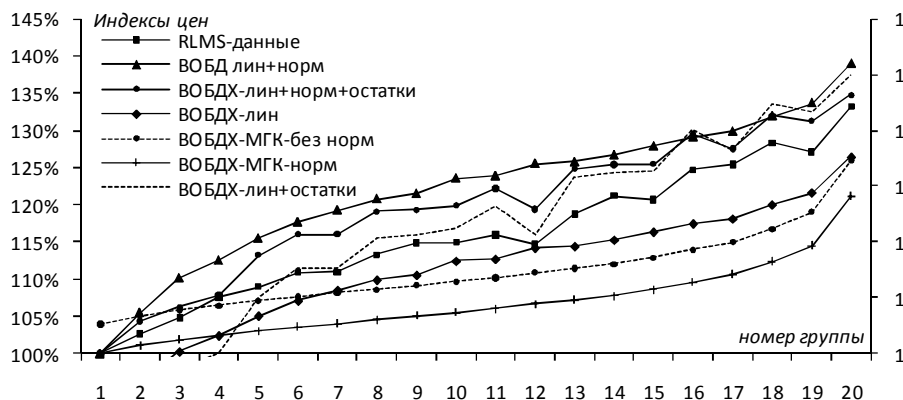


Рис. 5. Результаты регрессионного моделирования межгруппового индекса цен Фишера по данным RLMS для 20 групп

Подробнее рассмотрим предложенные корректировки и оптимальный набор факторов, включаемых в модель. Включение в модель регрессионных остатков уравнения, построенного по данным RLMS, при расчете индекса ВОБДХ, по-видимому, не имеет достаточных оснований. Несмотря на наличие малых колебаний в межгрупповой динамике индекса цен в выборке RLMS и всех остальных факторов в обеих выборках, нет априорных соображений, позволяющих предполагать согласованность таких колебаний в двух выборках. Разбиение каждой выборки на группы происходило независимо, и, вообще говоря, группы не соответствуют друг другу по абсолютным значениям параметров.

Нормировка индекса цен для первой группы к единице, по-видимому, не является принципиальным обстоятельством, а скорее техническим приемом, облегчающим интерпретацию получаемой динамики. Кроме того, она позволяет отказаться от распространения свободного члена линейной в логарифмах регрессии по данным RLMS на ВОБДХ и ограничиться только распространением коэффициентов при факторах, так как при такой нормировке и функциональной форме уравнения свободный член не влияет на окончательные значения нормированных рассчитанных индексов. В случае линейной модели можно предложить в качестве нормировки не деление на значение для первой группы, а вычитание его из рассчитанного индекса для каждой группы.

Заметим также, что нормировка рассчитанных индексов цен для ВОБДХ не позволяет напрямую сравнивать межгрупповую динамику с исходной для RLMS в силу несовпадения абсолютных значений характеристик групп в выборках. Для более корректного сопоставления нормировка должна проводиться не для группы с одним номером, а для групп с близкими значениями параметров.

Выбор факторов должен производиться с учетом природы рассматриваемых характеристик групп, а также принимая во внимание очень сильно выраженную мультиколлинеарность. В такой ситуации в уравнение не может быть включено много регрессоров. Различные комбинации включения факторов показали, что для 20 групп значимыми оказываются не больше двух. Окончательный набор включаемых в регрессию факторов определялся рядом критериев. Коэффициенты должны быть значимыми на приемлемом уровне (5–10%), знаки коэффициентов должны совпадать со знаками парных корреляций (требование гармонических регрессий) [4], регрессия должна иметь большой  $R^2$ , включенные в модель факторы должны достаточно точно моделировать удаленные.

Напомним, что группы выделялись исходя из показателя расходов на питание на одного члена домохозяйства. Так что информация об этом параметре неявно присутствует в самой структуре данных. Информация о составе домохозяйства представлена в обследованиях не полностью. Известно, что люди разного возраста и социально-демографических характеристик имеют разную структуру и объем потребления. Для учета такой дифференциации необходимо приведение размера домохозяйств к некоторым условным единицам, что сложно реализуется на данных рассматриваемых выборочных обследований. Кроме того, именно эти два фактора наилучшим образом объясняются остальными. Третьим, исключенным из модели регрессором стала доля расходов на питание – именно она оказывалась незначима как в линейной, так и в логлинейной моделях. Аналогичные расчеты проводились для случаев 20 и 50 групп. Результаты моделирования оказались близкими, и все качественные выводы сохраняются.

Таким образом, окончательно были построены две регрессии межгруппового индекса цен Фишера на общий уровень расходов на человека в группе и среднюю долю городского населения в группе. Для случая разбиения на 20 групп уравнения приняли вид

$$(6) \quad IPF = 0,82 + 0,31 \cdot Urban + 1,57 \cdot 10^{-5} \cdot Total\_exp, R^2 = 0,97, \\ (35,1) \quad (7,51) \quad (10,6)$$

$$(7) \quad \ln(IPF) = -1,07 + 0,05 \cdot \ln(Urban) + 0,14 \cdot \ln(Total\_exp), R^2 = 0,96. \\ (-7,53) \quad (1,72) \quad (9,42)$$

Далее в уравнения (6) и (7) были подставлены соответствующие значения для данных ВОБДХ и получены расчетные значения индексов цен. Для более удобного сравнения результатов нормировка проводилась не на значение индекса в первой группе, а соответствующее значение во второй, наиболее близкой к первой, группе по данным RLMS по расходам на питание на одного члена домохозяйства.

Из рис. 6 видно, что межгрупповая динамика восстановленных значений индекса цен по обеим моделям оказывается достаточно близкой к исходной динамике по данным RLMS. Существуют некоторые различия в расчетных значениях разных моделей – так, линейная модель предполагает более быстрый рост в середине диапазона и несколько более медленный в старших группах, как по сравнению с исходной динамикой, так и с логлинейной моделью. Логлинейная модель предлагает более плавную динамику, которая также ближе к исходной. Для логлинейной модели в условиях более традиционной нормировки – деление рассчитанных значений индексов на значение в группе, – которая принята за базовую, нет необходимости переноса

модельного значения свободного члена, что упрощает расчеты и облегчает интерпретацию.

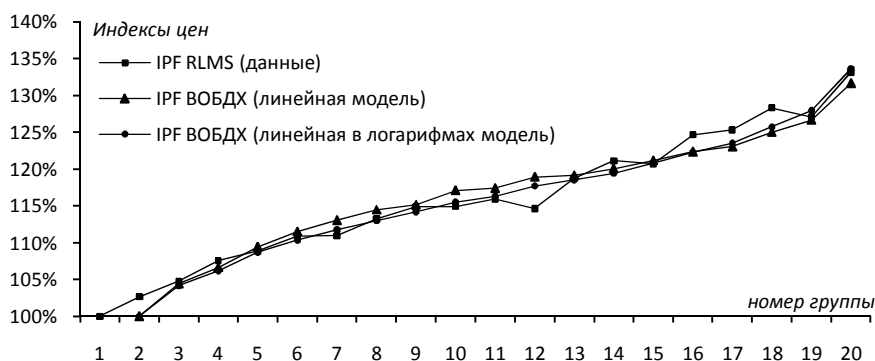


Рис. 6. Две регрессионные модели восстановления межгрупповых индексов цен Фишера для ВОБДХ с помощью выборки RLMS

Полученные с помощью такой процедуры индексы цен покупок продуктов питания для 20 или 50 групп домашних хозяйств, выделенных по уровню подушевых расходов на питание, могут быть использованы в дальнейшем моделировании структуры спроса домашних хозяйств на отдельные товары и услуги или группы продуктов. Очевидно, что аналогичным образом рассчитываются и другие индексы цен. Такая процедура легко обобщается на случай большего числа объясняемых параметров, если такая информация будет в обеих выборках, для которых производится «имплантация» индексов, или пересмотрена для другого способа выделения однородных групп домашних хозяйств.

Основным результатом является предложенный и реализованный механизм дополнения информации одного статистического источника недостающими данными из другого даже в случае, когда они не могут быть напрямую объединены.

## 5. Заключение

Полученные результаты свидетельствуют о том, что группы домашних хозяйств с разным уровнем доходов или объединенные по другому аналогичному признаку имеют разные потребительские предпочтения, в том числе в отношении не только структуры потребления, но и цен покупок. Такая дифференциация может быть следствием различного предпочтения качества приобретаемых товаров, которое не в полной мере отражено в данных микрообследований бюджетов и потребления, и не подтверждает гипотезу эндогенности средних цен покупок. Показано наличие зависимости межгрупповой динамики индекса цен покупок продуктов питания от микроэкономических характеристик домохозяйств, в том числе подушевой суммы расходов на питание.

В этой ситуации моделирование домашних хозяйств как единого агента для целей получения прогноза может предоставить только количественные данные на макроэкономическом уровне. Получить прогноз динамики структуры потребления по то-

варам возможно при учете внутренней структуры распределения потребителей в сочетании с ее прогнозом. Для корректного построения таких моделей необходимо максимально полно использовать всю доступную статистическую информацию. В том числе важной задачей является разработка механизма совместного использования информации из разных источников и дополнения их данными друг друга при частичном удовлетворении целям моделирования каждого из них в отдельности.

Такое моделирование представляется особенно актуальным для оценки эффекта и влияния различных мер по воздействию на спрос, в том числе тарифных политик. Очевидно, что глобальные изменения, хотя и не влияют напрямую на спрос населения на отдельные группы товаров и услуг, могут иметь значительные косвенные эффекты. Такие эффекты лишь частично улавливаются при моделировании потребления на макроэкономическом уровне. Так, например, сокращение реальных располагаемых доходов населения может быть крайне неравномерным и значительно изменить структуру спроса. Причем, как показал проведенный анализ, изменения в структуре могут не ограничиваться изменениями в соотношении потребления крупных агрегатов (питание, непродовольственные товары, услуги и т.д.). Различные группы населения могут предъявлять достаточно дифференцированный спрос и внутри отдельных групп продуктов, о чем, в частности, свидетельствует эндогенность цен покупок продуктов питания.

В текущей экономической ситуации принимается и будет приниматься целый ряд мер, оказывающих самое непосредственное влияние на спрос на товары и услуги со стороны населения. Так, во многих регионах происходило повышение тарифов ЖКХ, в стране пересматриваются акцизы и импортные пошлины на многие товары, изменяется курс рубля к валютам основных торговых партнеров – все это вызывает значительные изменения цен (в первую очередь относительных) на многие продукты. Для проведения наиболее верной государственной политики в этой области чрезвычайно важно максимально точно спрогнозировать реакцию спроса на такие воздействия.

Детальное исследование такой реакции должно проводиться в сочетании с моделированием эластичности спроса на отдельные продукты и группы. Подобное исследование является отдельной самостоятельной проблемой и требует глубокого системного подхода. В нашей работе не ставилась такая цель, но из полученных результатов следует, что, используя традиционные инструменты изучения реакции спроса на внешние воздействия, необходимо учитывать и динамическую структуру населения, в том числе распределение по доходам.

\* \*

\*

### СПИСОК ЛИТЕРАТУРЫ

1. Аллен Р. Экономические индексы. М.: Статистика, 1980.
2. Варшавский А.Е. Проблемы инноваций: риски и ответственность (на примере рынка продуктов питания и внутреннего потребления). М.: ЦЭМИ РАН, 2009.
3. Данные Российского мониторинга экономического положения и здоровья населения (RLMS). (<http://www.cpc.unc.edu/projects/rlms>).



4. *Ершов Э.Б.* Конкурирующие регрессии: критерии и процедуры отбора // *Экономический журнал ВШЭ*. 2008. Т. 12. № 4. С. 488–511.
5. *Ершов Э.Б., Матъцин М.С.* Экономическая теория и статистическая практика анализа потребительского поведения домашних хозяйств / Конспект доклада на Российском экономическом конгрессе. М.: ИЭ РАН, 2009. ISBN 987-5-9940-0219-3.
6. *Кёвеш П.* Теория и практика экономического анализа. М.: Финансы и статистика, 1990.
7. *Матъцин М.С.* Проблемы использования результатов выборочных обследований домашних хозяйств для моделирования структуры их расходов / Конспект доклада на Российском экономическом конгрессе. М.: ИЭ РАН, 2009. ISBN 987-5-9940-0219-3.
8. Микроданные Выборочного обследования бюджетов домашних хозяйств, проводимого Росстатом (ВОБДХ). (<http://www.micro-data.ru>).
9. *Пеникас Г.И.* Анализ эволюции потребительского поведения в России за период 2000–2005 гг. // *Экономический журнал ВШЭ*. 2008. Т. 12. № 4. С. 512–542.
10. Российский статистический ежегодник. М.: Росстат, 2008.
11. *Consumer Price Index Manual: Theory and Practice*. Geneva: International Labor Office, 2004. (Издание на русском языке: Вашингтон: Международный валютный фонд, 2007.)
12. *Deaton A., Muellbauer J.* *Economics and Consumer Behavior*. Cambridge, 1980.
13. *Durlauf S.N., Blume L.E.* (eds). *The New Palgrave Dictionary of Economics*. 2<sup>nd</sup> ed. Palgrave Macmillan, 2008.
14. *Feenstra R.C., Reinsdorf M.B.* An Exact Price Index for the Almost Ideal Demand System // *Economics Letters*. 2000. Vol. 66.
15. *Kalman P.J.* Theory of Choice Price Enter the Utility Function / Logistics Research Project. The George Washington University, Serial 13589, 4 August 1966.
16. *Kalman P.J.* Theory of Consumer Behavior When Prices Enter the Utility Function // *Econometrica*. 1968. Vol. 36. № 3–4.
17. *Marschak J.* Money Illusion and Demand Analysis // *Review of Economics and Statistics*. 1943. Vol. XXV. P. 40–48.
18. *Pesaran M., Schmidt H.P.* (eds.) *Handbook of Applied Econometrics*. Vol. I (Macroeconomics) and II (Microeconomics). Blackwell Publishing, 1999.
19. *Veblen T.* *The Theory of the Leisure Class*. N.Y.: Vanguard Press, 1912.