

Обследования экономически активного населения: занятость, безработица и неполная занятость

Методологическое руководство МБТ

- Концепции
- Методы

Р. Хуссманнс, Ф. Мехран, В. Верма

В 2 книгах

Книга 2



36107

Москва
ФИНСТАТИНФОРМ
1994

Оригинальное издание данной работы опубликовано МБТ (Женева) в 1990 г. под названием "Surveys of economically active population, employment, unemployment and underemployment".

Перевод и издание с разрешения МБТ.

Обозначения, используемые в публикациях МОТ и соответствующие практике ООН, а также подача материала в них не служат отражением точки зрения Международного бюро труда относительно статуса какой-либо страны, района или территории, либо их властей, равно как и относительно их границ.

За оценки, данные в подписанных статьях, исследованиях и других материалах, отвечают исключительно их авторы, а публикация не означает согласия Международного бюро труда с выраженными в них взглядами. Ссылка на наименования фирм и продукты или процессы не означает их признания Международным бюро труда, как и отсутствие упоминания какой-либо фирмы, продукта или процесса не свидетельствует о неодобрении их со стороны МБТ.

Перевод с английского *М.З.Штернгарца*

ЧАСТЬ ВТОРАЯ

МЕТОДЫ

Глава 10. ПЛАНИРОВАНИЕ, ОРГАНИЗАЦИЯ И СОВЕРШЕНСТВОВАНИЕ ОБСЛЕДОВАНИЙ

1. ВВЕДЕНИЕ

Часть 2 настоящего руководства посвящена различным методологическим вопросам организации и проведения обследований экономически активного населения, причем особое внимание уделено условиям и требованиям осуществления этой работы в статистически менее развитых странах. В настоящей главе содержится обзор аспектов планирования и организации обследований рабочей силы или, в более общем виде, обследований экономически активного населения. Более полное и детальное рассмотрение вопросов, связанных с организацией и проведением обследования, содержится в последующих главах.

Во многих странах, и особенно в статистически менее развитых, выборочные обследования домашних хозяйств являются главным источником информации об экономически активном населении. Такова основная причина, по которой этому виду обследований уделено главное внимание в настоящем руководстве. Но выборочные обследования домашних хозяйств не служат единственным источником такой информации. В парагр. 2 настоящей главы рассмотрены преимущества и недостатки других источников, например, переписей населения и учета в административных органах. Различные источники можно сочетать и использовать так, чтобы они дополняли друг друга в качестве источников требуемой информации. Выбор наиболее удачного сочетания источников данных и является первым шагом в статистическом планировании.

При обследованиях домашних хозяйств, направленных на получение информации о занятости, безработице, неполной занятости и других аспектах экономической активности, возможны различные организационные варианты. Определяющим моментом структуры обследования служит прежде всего *характер требуемых данных*. Обследования можно проводить на непрерывной основе для получения динамических рядов текущих данных или менее часто с целью получить информацию скорее о структурных сдвигах и долговременных тенденциях. Важным фактором может быть необходимость увязки с другими обследованиями, так как обследования домашних хозяйств все чаще становятся источником данных о многих других явлениях, помимо рабочей силы. Эти и другие факторы определяют общую структуру обследования, о чем говорится в парагр. 3 настоящей главы.

Параграф 4 посвящен основным вопросам планирования и организации обследований экономически активного населения. Они включают определение структуры и масштабов обследования, планирование и организацию работ и другие аспекты с учетом трех групп факторов: цели обследования и требования к данным; практические ограничения и условия проведения обследования; требование увязки с другими статистическими мероприятиями, в частности, с другими видами обследования домашних хозяйств.

Многие страны уже проводят обследования рабочей силы, и задача состоит в их периодическом переосмыслении и улучшении существующей системы обследований. Хотя при совершенствовании обследований применимы общие принципы их организации, при пересмотре действующей системы необходимо принимать во внимание много дополнительных факторов. Поэтому настоящая глава заканчивается парагр. 5 о целях и стратегии совершенствования обследований.

2. ИСТОЧНИКИ ДАННЫХ ОБ ЭКОНОМИЧЕСКИ АКТИВНОМ НАСЕЛЕНИИ

Переписи населения и выборочные обследования домашних хозяйств и индивидов обычно дают всесторонние данные об экономически активном населении, которые можно увязать с информацией по смежным вопросам. Обследования предприятий и учетные данные административных органов также служат источниками получения в ряде случаев более точной, оперативной и подробной информации об отдельных составляющих экономически активного населения. Различные источники информации следует рассматривать как взаимодополняющие, и их надо использовать, когда это необходимо, для получения интегрированных данных (МКСТ. Резолюция I, парагр. 3 (ILO, 1983)).

Выборочные обследования домашних хозяйств

Обследования домашних хозяйств позволяют одновременно измерить число занятых, безработных и экономически неактивных. Их можно организовать таким образом, чтобы охватить практически все население страны, все сферы экономической деятельности, все отрасли экономики и все категории трудящихся, включая владельцев собственных предприятий и неоплачиваемых работающих членов их семьи, а также лиц, занятых случайной и временной работой. Это обеспечивает уникальную возможность получить информацию о всей рабочей силе и ее структуре. Понятия, определения и степень детализации вопросов можно легко приспособить к требованиям, предъявляемым к информации, поэтому можно выявить различные формы участия в рабочей силе отдельных групп населения. Возможна при этом также значительная гибкость в отношении охвата вопросов данным обследованием. А так как происходит непосредственный контакт с домашними хозяйствами и индивидами, в ходе обследования, кроме данных о характеристиках рабочей силы, можно при сравнительно низких дополнительных затратах получить еще информацию о демографических и социально-экономических характеристиках индивидов и домашних хозяйств (United Nations, 1984, paras. 11.10 — 11.11). Это открывает много возможностей для анализа данных. При должной их организации обследования домашних хозяйств могут также обеспечить получение информации о семейных и других малых предприятиях.

Обследования домашних хозяйств служат главным источником информации об экономически активном населении, однако некоторые данные можно получить из других источников. Важнейшее решение при статистическом планировании связано с выбором правильной комбинации методов, позволяющих наилучшим образом получить данные в конкретной ситуации. Помимо обследований домашних хозяйств основные источники данных о рабочей силе включают: 1) переписи населения; 2) переписи и выборочные обследования предприятий; 3) различные учетные записи административных органов. Названные источники различаются охватом, масштабами, единицами измерения и методами сбора данных. Каждый источник обладает преимуществами и недостатками с точки зрения издержек, качества и характера выдаваемой информации. Как правило, один подход предпочтительнее другого в некотором отношении, и наоборот. Поэтому различные источники скорее дополняют друг друга, чем конкурируют между собой и

взаимно исключают друг друга. Их результаты можно сочетать при условии, что понятия, определения, охват, отчетный период, классификации и т.п. согласуются между собой, насколько это возможно. В последующей части настоящего раздела рассмотрены альтернативные источники информации об экономически активном населении в сопоставлении с выборочными обследованиями домашних хозяйств.

Переписи населения

Переписи населения и обследования домашних хозяйств в принципе охватывают одно и то же население и используют одни и те же единицы измерения (домашние хозяйства и индивиды). Различие в использовании переписей населения и обследований домашних хозяйств связаны главным образом с различиями в масштабах операции (полный охват в отличие от выборки), а это влечет за собой различия в методологии, практике проведения, временных рамках и сложности сбора данных. Главная цель переписи — получить относительно краткую, но всеохватывающую информацию о размерах и основных характеристиках населения, а также дать максимально детализированные сведения по регионам или малым населенным пунктам, или по отдельным группам населения. В рамках переписи населения обычно изучают лишь немногие аспекты экономической характеристики. Согласно Рекомендациям ООН, включению подлежат следующие аспекты, причем первый набор считается приоритетным (Unitea Nations, 1980, paras. 2.180—2.219):

статус экономической активности, занятие, отрасль, статус занятости;

отработанное время, доход, сектор занятости.

Число вопросов, используемых в переписи для изучения перечисленных моментов, обычно ограничено — по одному вопросу для каждого пункта. В отличие от этого, выборочное обследование домашних хозяйств, благодаря его меньшему масштабу, можно использовать для получения широкого круга данных для различных видов анализа. Его можно более гибко приспособить к различным потребностям пользователей и методам сбора данных. Хотя обследования домашних хозяйств требуют немалых затрат, они явно обходятся дешевле, нежели полные переписи. Их можно повторять более часто, обеспечивая тем самым информацию об изменениях за некоторый период времени. Благодаря меньшему объему выборочные обследования позволяют лучше контролировать ответы и избе-

гать ошибок, не связанных с выборкой, а их результаты можно обработать быстрее.

В связи с ограниченным размером выборок основной недостаток обследований домашних хозяйств заключается в их неспособности обеспечить достаточно подробные данные по небольшим районам и подгруппам населения. Кроме того, выборки умеренных размеров часто обеспечивают удовлетворительные оценки пропорций, уровней, коэффициентов и т.д., но они меньше подходят для выявления суммарных показателей (например, общего числа безработных в данной группе населения) и их изменений — т.е. моментов, которые могут представлять особый интерес для пользователя. Для получения надежных оценок таких агрегированных величин обычно необходимо дополнить или адаптировать выборочные данные, прибегая к информации из других источников, нередко из переписей населения или списков жителей, где они имеются.

Взаимодополняющие функции и совместное использование переписей населения и выборочных обследований охарактеризованы во многих источниках (United Nations, 1984, paras. 1.7 — 1.11; Kish and Verma, 1986). Выборочный метод можно с пользой применить для ускорения планирования, проверки, контроля, оценки, обработки и дополнения сбора данных переписи. В свою очередь, переписи обеспечивают инфраструктуру, основу для выборки, исходные данные и т.п., что необходимо для проведения выборочных обследований домашних хозяйств, а также дают общий толчок развитию статистического дела. Указанные моменты особенно важны при проведении крупномасштабных выборочных обследований с широким охватом, например, общенациональных обследований рабочей силы. Встречаются также сочетания обоих подходов. Во многих странах переписи включают два элемента: а) учет населения и его основных графических и смежных характеристик на 100%-ной основе, дополненный б) крупной выборкой, увязанной с переписью и охватывающей более широкий круг вопросов. Такая схема способна существенно повысить роль переписи как источника данных об экономически активном населении и по смежным вопросам.

Следует также отметить, что данные переписей населения и обследований рабочей силы все чаще используют в сочетании с соответствующими статистическими методами для получения текущих оценок для небольших районов и динамики показателей после переписи.

Переписи и выборочные обследования предприятий

В зависимости от конкретных требований данные о занятости и по смежным вопросам можно также получить на основе переписей и выборочных обследований предприятий в отличие от переписей населения и выборочных обследований домашних хозяйств. Различие между двумя подходами заключается в охвате населения, единице измерения и объеме получаемой информации.

По сравнению с обследованиями домашних хозяйств, переписи и обследования предприятий могут иметь более конкретную направленность, а следовательно, быть более избирательными и экономными, но в то же время они более ограничены по охвату и содержанию. При отборе объектов обследования различают два типа предприятий: крупные предприятия, принадлежащие к более организованному сектору экономики; как правило, они так или иначе зарегистрированы, а число занятых в них превышает определенный минимум; небольшие предприятия в относительно неорганизованном секторе, это могут быть несемейные и чисто семейные предприятия, принадлежащие домашним хозяйствам или нескольким компаньонам. Принимая во внимание большие расходы и технические сложности, а также цели обследований, обычно ими охватывают лишь крупные предприятия. Поэтому полученные в их результате данные о занятости чаще всего относятся к наемным работникам в организованном секторе экономики. Такие обследования нередко охватывают лишь отдельные сферы экономики, например, горнодобывающую промышленность, обрабатывающую промышленность или строительство. Обычно они базируются на имеющихся перечнях предприятий, и им в той или иной мере присущи недостатки, свойственные таким перечням.

Кроме того, правила соответствия между различными типами единиц обследования могут быть весьма сложными. Может оказаться затруднительным выявление и взаимное согласование единиц, используемых для выборки, с единицами, представляющими информацию, и единицами счета и анализа, по которым желательно получить информацию. Общеизвестны практические трудности, связанные с последовательным разграничением предприятий и их подразделений. Иллюстрацией иного рода может служить то, как при обследовании предприятий можно столкнуться с тем, что в ведомостях на получение заработной платы не значатся работающие, но временно отсутствующие трудящиеся, и в то же время по несколько раз фигурируют совместители. Такого рода проблемы при обследовании домашних хозяйств не возникают.

Достоинство обследований предприятий заключается в их большей *избирательности*, как по охвату, так и по содержанию. Когда интерес представляют конкретные отрасли, обследования предприятий при наличии надежных их перечней способны обеспечить более удачные выборки и процедуры по сравнению с обследованиями домашних хозяйств, охватывающими все население. При обследованиях предприятий можно получить более достоверную и подробную информацию, особенно когда они опираются на ведомости заработной платы и другие доступные учетные данные. Обследования этого рода могут позволить собрать информацию о многих других экономических показателях и явлениях — продукции, издержках, инвестициях, технологических и организационных факторах, которую затем можно непосредственно соотнести с информацией о занятости, заработной плате и производительности и которая может послужить более широкой основой для анализа экономической активности. А данные о занятости можно более точно соотнести с данными о заработках, квалификации, занятии и отрасли.

Обследования предприятий могут быть также более экономичными и оперативными по сравнению с обследованиями домашних хозяйств. Связано это с тем, что респонденты при обследовании предприятий расположены более компактно и с ними легче осуществить контакты. Даже в развивающихся странах вместо более дорогостоящего личного опроса можно прибегнуть к почтовой анкете или телефонному опросу. Еще один фактор, способствующий снижению затрат, заключается в том, что требуемую информацию о всех занятых на крупном предприятии можно получить от одного или немногих респондентов.

Но если обратиться к небольшим предприятиям, различие между ними и домашними хозяйствами не столь очевидно. Перечни небольших предприятий, как правило, отсутствуют, так как их текучесть велика и они трудноразличимы, поэтому единственно целесообразный подход, как и при обследовании домашних хозяйств, сводится к многоступенчатой выборке. Но, в отличие от домашних хозяйств, даже небольшие предприятия часто распределены по территории страны неравномерно и характеризуются высокой степенью концентрации по видам экономической деятельности. С целью добиться большей экономии и эффективности в организации обследования часто необходимо прибегать к информации о размещении предприятий, имеющейся в результатах переписей и в других источниках (Murthy and Roy, 1970).

Административные источники

Данные, основанные на административном учете, обычно представляют собой побочный продукт деятельности административных органов. Поэтому они могут представлять весьма экономичный источник статистической информации. Нередко записи производят на непрерывной основе, отсюда они могут служить полезным источником данных в динамике и за продолжительный период. Но им также могут быть присущи серьезные недостатки, например, ограниченный охват и содержание, негибкие понятия и определения, неполнота, непоследовательность, а также ограниченная доступность из-за юридических и административных моментов. Как отмечается в "Справочнике по обследованиям домашних хозяйств" (ООН, 1984, парагр. 1.13, 1.14), в развивающихся странах с неорганизованными рынками труда административные источники типа данных органов страхования на случай безработицы и биржи труда часто вообще отсутствуют или ограничены узким кругом трудящихся. Разумеется, когда данные существуют и обобщаются часто и через регулярные интервалы, их можно с пользой применить для некоторых видов анализа. Для получения информации о занятости в государственном секторе целесообразно также воспользоваться другими данными административного учета, например, ведомостями заработной платы и карточками правительственных организаций, государственных предприятий и других учреждений. К такого рода источникам чаще прибегают, однако, в развитых странах.

Административные источники можно использовать при составлении и обновлении перечней для обследования занятости и аналогичных работ, базирующихся на выборке предприятий, в которую попадают в основном относительно крупные предприятия в организованном секторе экономики.

3. Структура и организация обследований

Согласно резолюции, принятой на 13-й МКСТ, программа статистического изучения экономически активного населения должна охватывать все сферы экономической деятельности, все вектора экономики и все категории трудящихся, а также по возможности полнее гармонизировать с другими видами экономической и социальной статистики. Как вытекает из предыдущего параграфа настоящей главы, обследования домашних хозяйств наиболее полно отвечают названным требованиям. 13-я МКСТ далее определила:

“Программа (статистического изучения экономически активного населения) должна, в частности, учитывать как краткосрочные, так и долгосрочные потребности, то есть предусматривать частое получение данных на регулярной основе для текущих нужд и сбор информации через более продолжительные промежутки времени для более углубленного структурного анализа и формирования базы данных.

А. Программа текущих обследований должна охватывать данные об экономически активном населении и его составе таким образом, чтобы можно было правильно выявить тенденции и сезонные колебания.

Б. Программа менее частых статистических работ может включать переписи и обследования, которые призваны обеспечить получение: 1) всесторонних данных об экономически активном населении; 2) более углубленных данных о характере активности этого населения в течение года и взаимосвязи между занятостью, доходом и другими социальными и экономическими характеристиками; 3) данных по другим частным вопросам (например, дети и молодежь, женские домашние хозяйства), которые вытекают из долгосрочных постоянных потребностей” (13-я МКСТ, 1982, Резолюция I, парагр. 2) (ILO, 1983).

Минимальные требования, относящиеся к частоте получения, содержанию и классификации данных, содержатся в Рекомендациях по статистике труда (№ 170) принятой Международной конференцией труда в 1985 г. в качестве дополнения к Конвенции о статистике труда (№ 160).

1. Текущие данные об экономически активном населении, занятости, где это необходимо — о безработице и, где это возможно, о видимой неполной занятости следует собирать по меньшей мере раз в год. Эти данные следует классифицировать по признаку пола и, где это возможно, по возрастным группам и отраслям экономической деятельности.

2. Учитывая долговременные потребности в подробном анализе и базисных показателях, данные о структуре и распределении экономически активного населения необходимо собирать по меньшей мере раз в 10 лет. Указанные данные следует классифицировать по меньшей мере по половозрастным группам, профессиональным группам или уровням квалификации, отраслям экономической деятельности, географическим районам и статусу занятого (например, работодатель, самостоятельный работник, работающий по найму, неоплачиваемый трудящийся семейного предприятия, член производственного кооператива).

Обследования домашнего хозяйства для учета экономически активного населения можно организовать и провести по-разному. Главными определяющими факторами являются *содержательные цели*, т.е. содержание, сложность и периодичность требуемой информации. Обследование можно организовать таким образом, чтобы регулярно получать динамические ряды данных об уровнях и тенденциях занятости; альтернативно или в дополнение к этим данным обследование может сосредоточиться на менее частой информации структурного характера, представляющей долговременный интерес. Данные можно детализировать таким образом, чтобы получить информацию, например, о динамике рабочей силы или переливах между отдельными категориями рабочей силы; или же, наоборот, обследование можно ограничить немногими основными характеристиками рабочей силы, например, уровнем занятости и безработицы. Такого рода содержательные соображения определяют также время, частоту и период проведения обследования, характер выборки и другие аспекты работы. Другим важным фактором, определяющим структуру и организацию обследования, может служить необходимость увязки с иными обследованиями как в части содержания, так и практического проведения работ.

Непрерывные и случайные обследования

Непрерывные обследования для получения текущих данных. Непрерывные обследования проводят главным образом для получения динамических рядов данных о текущих уровнях и тенденциях; при должной организации обследования и достаточно высоком качестве данных можно получить оценки общих изменений и сдвигов в структуре занятости и видах экономической деятельности. Как правило, такое обследование включает продолжающуюся серию периодов ("раундов"), каждый из которых предназначен дать отдельные оценки, относящиеся к определенному периоду времени.

К непрерывным обследованиям прибегают для слежения за функционированием экономики; получения показателей изменений в текущем уровне участия в рабочей силе, занятости, безработице и неполной занятости; а также для измерения тенденций, циклических и сезонных колебаний указанных уровней. Хотя при соответствующей его организации обследование может также дать информацию об общих изменениях и сдвигах (переливах), сделать это с достаточной надежностью удается не всегда.

Обычно различают два вида непрерывных обследований. Первая форма — это проведение обследования на *постоянной основе*, при

этом обследовании сбор информации на местах производят непрерывно. Делают это, как правило, на основе скользящего отчетного периода, т.е. периода, относящегося к определенному отрезку времени, непосредственно предшествующему опросу, причем этот период различен в зависимости от даты опроса. Если период обследования достаточно продолжителен, его можно разделить на подпериоды, каждый из которых охватывает представительную выборку в течение части общего периода проведения обследования. Деление на подпериоды позволяет лучше, в смысле большей представительности, охватить выборку за более краткие отрезки периода обследования (например, месяцев или кварталов года). Полученные по подпериодам результаты можно поэтому использовать для изучения сезонных и иных колебаний в течение года. Система обеспечивает также более частую и оперативную публикацию результатов. Различие между периодами и подпериодами в основном сводится к размеру выборки. По результатам каждого периода с выборкой более крупного размера, полученной на основе сложения данных по подпериодам, можно составить более подробные таблицы и осуществить более детальный анализ; тогда как по отдельным подпериодам размер выборки может обеспечить лишь некоторые оценки достаточной точности без их дальнейшего деления по районам и т.д. Когда получение данных распределяют по представительным подвыборкам для каждого отрезка времени (подпериода), улучшаются агрегированные результаты для всего периода, так как учитываются сезонные и другие колебания в течение периода или они сглаживаются более сбалансированным образом. Деление на подпериоды может также иметь существенные организационные преимущества: сбор данных можно лучше проконтролировать и более равномерно распределить во времени. Разумеется, транспортные расходы при этом возрастают, так как в ходе каждого подпериода необходимо охватить территориально достаточно разбросанные выборочные единицы.

Второй вид — *периодические* обследования с прерывистым сбором данных в течение относительно кратких интервалов времени. Такой метод может обладать определенными преимуществами. Во-первых, он способен облегчить контроль и проведение сбора данных. Во-вторых, становится легче получать информацию при “фиксированном” отчетном периоде, т.е. одинаковом отчетном периоде для всех респондентов в виде фиксированных календарных дат. Как указано ниже, для некоторых целей необходимо и/или предпочтительно использовать именно фиксированный период.

Однако такая организация работы имеет некоторые недостатки. Во-первых, средние условия в течение, например, года путем периодических обследований выявить труднее, нежели с помощью непрерывных обследований, когда сбор данных распределен равномерно по всему периоду. Во-вторых, периодические обследования не позволяют равномерно загрузить счетчиков постоянной работой. Поэтому необходимо сделать выбор между: а) использованием постоянных счетчиков с предоставлением им свободного времени в течение части года; б) наймом временных штатов во время сбора данных; в) наличием постоянных штатов с использованием работников вне периодов сбора данных для других работ, например, обновления выборки, редактирования, кодирования и т.д. Первый вариант (а) явно расточителен, хотя он серьезно рассматривался в определенных ситуациях, когда использование хорошо подготовленных счетчиков считали важным для получения качественных данных, а расходы были относительно невелики из-за низкой заработной платы. Применительно к варианту (б), использованию временных работников, был предпринят анализ ряда национальных обследований рабочей силы (ILO, 1986), который показал, что в половине стран, где проводят обследования на периодической (в отличие от непрерывной) основе, по меньшей мере отчасти прибегали к найму некоторых временных работников, чтобы решить проблему неравномерной загрузки, присущей обследованиям. Напротив, согласно тому же источнику, практически ни в одной из стран с непрерывным обследованием к услугам временных работников не прибегали. Разумеется, укомплектование кадрами зависит от множества факторов, однако колебания объема работ — это, очевидно, важное соображение.

Вариант (в) распространен широко, особенно в странах со сравнительно хорошо сложившейся статистической инфраструктурой, когда постоянные работники, расположенные в различных частях страны, способны осуществить различные работы по сбору данных, включая опросы в рамках обследований домашних хозяйств. В перерывах между сбором данных для этих обследований такие работники могут продолжать сбор данных другого рода. Возможен альтернативный подход, когда этот персонал используют между раундами обследования не для сбора других данных, а для редактирования, кодирования, обобщения и оценки данных обследования рабочей силы, полученных в ходе предыдущего раунда работ. По мере развития микрокомпьютерной технологии становится возможной децентрализация большого числа функций обработки данных, например, их ввода и проверки. Участие счетчиков в подготовке и

обработке данных способно ускорить всю операцию. Это может послужить важным обстоятельством, так как обработка данных часто является узким местом, которое задерживает завершение обследований. Участие в обработке может также улучшить качество собираемых данных при последующих раундах, ибо оно способствует лучшему пониманию анкеты интервьюером.

Учитывая, что основная цель непрерывных обследований, будь то на постоянной или периодической основе, заключается в регулярном получении данных, совершенно необходимо, чтобы и результаты их сообщали регулярно и своевременно, а объем и сложность собранной информации соответствовали возможности организации их обработки и обнародования. В противном случае необработанные данные будут накапливаться во все возрастающих объемах, а цель обеспечения текущей информацией будет дискредитирована.

Много примеров непрерывных обследований рабочей силы можно встретить как в развивающихся, так и в развитых странах; анализ имеющегося в странах опыта содержится в работе МБТ (ILO, 1986).

Случайные обследования для получения информации структурного характера. Всесторонние обследования экономически активного населения можно проводить реже для получения исходных данных и подробной структурной информации. Она может включать, например, информацию об экономически активном населении в разрезе отраслей, занятий, статуса, о характере активности в течение года, стаже работы, совместительстве, образовании и профессиональной подготовке, отработанном времени, доходе от занятости и т.д. Аналогичным образом экономически неактивное население можно сгруппировать по видам и в зависимости от социально-экономических и демографических характеристик. При планировании развития страны такие обследования необходимы для анализа состояния в области занятости в начале планового периода и установления целей и заданий. Обследования такого типа не предназначены для получения непрерывного потока текущих данных или информации о краткосрочных переменах, но они очень удобны для получения информации, которая не требуется столь часто и характеризует существенные структурные характеристики и долгосрочные сдвиги. А так как эти характеристики быстро не меняются, не обязательно проводить такие обследования чаще одного раза в несколько лет. Например, их можно проводить раз в пять лет в виде обследований после переписей и между переписями в странах, где переписи проводят раз в десять лет. Во всяком случае, чаще проводить такие обследования просто нецелесообразно по экономическим соображениям.

Необходимо тщательно определить время проведения обследования. Результаты должны быть готовы тогда, когда структурные и исходные данные более всего нужны, например, при разработке планов развития. Для того чтобы эти результаты сохраняли ценность длительное время, проведение обследования не следует приурочивать к ненормальным и преходящим ситуациям в сфере занятости (если только анализ таких ненормальных ситуаций не является главной целью специального обследования). При выборе периода обследования надо учитывать также сезонные и иные краткосрочные колебания. По названным причинам может оказаться целесообразным растянуть сбор данных на весь год, чтобы охватить все сезоны; это можно осуществить на постоянной основе с непрерывным сбором данных в течение года (как это, например, делают в рамках раундов национального выборочного обследования рабочей силы в Индии) или в виде более сжатых периодов сбора данных на протяжении года.

Время от времени можно прибегать к специализированным обследованиям с целью более глубокого изучения некоторых взаимосвязей, особых явлений, проблем и вопросов, а также групп населения, представляющих специальный интерес. Возможными примерами являются взаимосвязи между затратами труда, профессиональной подготовкой и опытом работы, с одной стороны, и доходом от занятости, семейным доходом, благосостоянием и т.д., с другой. Однако такие обследования, как правило, достаточно сложны по содержанию, требуют специальной организации, специализированного персонала и сравнительно больших затрат и усилий. Поэтому их обычно проводят в разовом порядке или редко, охватывают они не всю страну и предполагают выборки меньших объемов и не очень разбросанные. Поскольку специализированные обследования «научно ориентированы» (т.е. направлены на получение информации, представляющей долговременный интерес с точки зрения лучшего понимания проблем, а не на решение сиюминутных задач), возможна значительная гибкость в выборе времени их проведения.

Наконец, можно отметить, что цели обследования иногда предполагают получение дополнительной информации о некоторых группах населения, представляющих особый интерес, например, инвалидах, мигрантах, женщинах — главах домашних хозяйств, безработной молодежи или неполностью занятых трудящихся. Если эти группы малочисленны, могут потребоваться специальные подходы, например, многофазный отбор с отсеиванием, с тем чтобы охватить достаточное число респондентов. При непрерывных обследованиях также существует возможность накопления таких случаев на основе нескольких раундов. Когда же требуемая дополнитель-

ная информация о специальных группах населения слишком подробна или сложна, может оказаться необходимым организовать ее сбор в виде операции, отдельной от основного обследования рабочей силы (с помощью, например, особых анкет и/или особых интервьюеров), но скоординированной с ним должным образом.

· Отчетный период

Еще один существенный элемент структуры обследования — это характер отчетного периода (или отчетных периодов) для сбора информации. В части I настоящего руководства уже были рассмотрены различные соображения, связанные с выбором отчетного периода. Здесь же мы обобщим основные моменты выбора периода в контексте различных видов обследований рабочей силы, о которых говорилось выше.

Напомним, что для измерения активности в данный момент необходим отчетный период в одну неделю или один день. Для измерения обычной активности используют продолжительный отчетный период, охватывающий целый год. Краткий отчетный период обычно подходит для обследований непрерывного типа, направленных в основном на получение текущих показателей. При менее частых обследованиях с целью выявления структурных характеристик, представляющих долговременный интерес, можно рассмотреть сочетание продолжительного и краткого отчетных периодов. Продолжительный отчетный период и выявление обычного статуса индивидов могут быть особенно уместны для исследований, призванных проанализировать сложные глубинные взаимосвязи между экономической активностью и другими показателями.

Отчетный период может быть фиксированным, в виде конкретного календарного года, или скользящим, в виде определенного отрезка времени, предшествующего моменту опроса. Каждый вариант имеет свои преимущества. *Фиксированный отчетный период* дает информацию, относящуюся к определенному времени, которое одинаково для всех обследуемых единиц. Это позволяет избежать “влияний периода” (т.е. особенностей уик-энда или конца месяца) на ответы различных респондентов. Иногда такой подход облегчает также увязку с данными из других источников, он может быть необходим с точки зрения некоторых юридических норм или традиций, требующих получения оценок для определенных (календарных) периодов. С другой стороны, при использовании фиксированного отчетного периода, как правило, период вспоминания (т.е. продолжительность времени, которое респондент вынужден вспо-

минать для ответа) не только удлиняется, но и может оказаться различным для разных респондентов в зависимости от времени проведения опроса данного домашнего хозяйства. Это может составлять серьезную проблему в случае, когда сбор данных относительно растянут во времени. Чтобы справиться с указанной проблемой, сбор данных следует проводить в сжатые сроки, т.е. на периодической, а не постоянной основе, но это в свою очередь приводит к неравномерной и нерегулярной загрузке персонала. Поэтому при непрерывных обследованиях, проводимых в виде периодических раундов, как правило, используют фиксированный отчетный период; так же поступают при случайных обследованиях, когда сбор данных осуществляют в течение короткого отрезка времени. Недостатки фиксированного отчетного периода, например, удлинение и неравномерность периода воспоминания, можно уменьшить, если сбор данных проводить в сжатые сроки. Проблемы неравномерной загрузки становятся менее острыми при увеличении частоты обследований. Так, ежемесячные обследования рабочей силы часто являются периодическими по своей природе (сбор данных осуществляется в течение одной или двух недель месяца), и в качестве фиксированного отчетного периода в них используют последнюю календарную неделю или календарный месяц (ILO, 1986). При менее частых периодических обследованиях охотнее прибегают к скользящим периодам.

При непрерывном сборе данных *скользящий отчетный период*, как правило, более предпочтителен. Сказанное справедливо и тогда, когда само обследование непрерывно и когда оно носит случайный “структурный” характер: в обоих случаях недостатки длительного и неравномерного периода воспоминания, присущие фиксированному отчетному периоду, усугубляются, так как сбор данных растягивается до года. Преимуществами скользящего отчетного периода являются более короткий период воспоминания и сглаживание влияния сезонных и аналогичных факторов в рамках всей выборки. Но результаты в целом относятся к менее четко определенному периоду времени, а для отдельных респондентов влияние различных сезонных и других подобных факторов может быть ощутимым.

Согласование с другими источниками

Во многих странах обследования домашних хозяйств все шире используются в качестве источника разнообразной статистической информации. В результате часто необходимо либо проводить обследо-

дования (экономически активного населения или иного характера) в виде части системы или программы единых обследований, либо по меньшей мере обеспечивать их тесную координацию. Указанные требования особенно важны применительно к обследованиям экономически активного населения, которые, как правило, являются всеохватывающими, общенациональными и сравнительно крупными по своим масштабам. Потребность увязки становится еще больше, когда обследования носят периодический или непрерывный характер, так как при этом могут пересекаться различные работы, одновременно проводимые статистическим органом.

Согласование обследований включает два широких аспекта:

а) координацию на организационном уровне для выработки единых процедур и подходов с целью добиться большей экономии и гибкости при проведении работ;

б) интеграцию на содержательном уровне для согласования вопросов и тем обследований с целью получения взаимно увязанных данных и их совместного анализа.

Координация означает, что отдельные обследования планируют и осуществляют во взаимосвязи на основе единых методов и инфраструктур, включая организационные рамки, условия выборки, технический, контрольный и конторский персонал, а также транспорт, обработку данных, печатание результатов, применение технических средств. Степень координации и совместного использования средств может быть разной в зависимости от вида задействованной организации, характера ее мероприятий, специальных требований, финансирования и других условий проведения обследований. Небольшие одноразовые обследования часто можно осуществлять на основе специальных методов и форм, но это не всегда возможно применительно к крупным начинаниям типа общенациональных обследований рабочей силы. Такие обследования нередко проводят в рамках регулярных операций национального органа статистики, и их планирование предполагает тщательное согласование с другими мероприятиями, оценку возможных ограничений и проблем (например, возросшую нагрузку на средства обработки данных и технический персонал), а также возможностей для увеличения гибкости при таком согласовании.

Интеграция на содержательном уровне предполагает исследование того же населения, использование одинаковых понятий и определений, единую систему классификации, а возможно, и стандартные вопросы обследования применительно к таким определениям, как возраст, пол, этническая группа, образование и статус в области активности, а также общие или пересекающиеся выборки

респондентов (United Nations, 1986, p. 6; 1983, para. 1364). Иногда для обозначения охвата многих тем в ходе одного обследования, применительно к единой выборке, а возможно, и при одном опросе респондента употребляют термин “полная интеграция”. Термин же “частичная интеграция”, напротив, характеризует ситуацию, когда для различных тем изучают одну и ту же выборку районов, но разные выборки домашних хозяйств в каждом районе. Лишь полная интеграция обеспечивает увязку данных на микроуровне¹.

На практике можно встретить различные формы и методы увязки обследований рабочей силы с другими обследованиями:

1) обследование рабочей силы можно организовать как операцию, более или менее самостоятельную по отношению к другим обследованиям;

2) обследование рабочей силы ограниченного содержания можно включить в какое-либо другое обследование в виде “модуля”, или наоборот;

3) более исчерпывающее обследование рабочей силы можно использовать для охвата смежных тем;

4) обследование рабочей силы может составлять часть многоцелевого обследования, охватывающего различные темы;

5) его можно проводить в качестве раунда продолжающейся системы обследований, при которой на каждом раунде изучают разные вопросы. Перечисленные возможные варианты рассмотрены ниже.

Самостоятельные обследования рабочей силы

В ряде развивающихся и развитых стран (например, Дании, Египте, Испании, Таиланде, Финляндии) проводят обследования, которые преимущественно или исключительно посвящены проблематике рабочей силы. В настоящем контексте их лучше всего именовать “самостоятельными” обследованиями рабочей силы. Их ориентация на единственную тему не исключает, разумеется, рабочей координации и использования совместных технических средств и общих подходов с другими обследованиями, равно как и общего охвата, одинаковых понятий, определений и классификаций. Под “самостоятельными” подразумевается лишь то, что они нацелены

¹ Термином “интеграция” иногда обозначают согласование содержательных и организационных аспектов обследований; термин “координация” будет в этом случае относиться к менее тесной форме интеграции. Однако предпочтительнее употреблять оба термина так, как это сделано выше: и координация, и интеграция могут быть различной степени тесноты.

на одну проблему и в значительной мере автономны с точки зрения планирования и проведения. Такая самостоятельность иногда обеспечивает лучший контроль и наблюдение, а также большую гибкость в планировании и осуществлении обследования.

"Модуль" рабочей силы, включенный в другие обследования

Важную информацию об экономических характеристиках и деятельности населения часто собирают в ходе переписей населения, демографических обзоров, обследований бюджетов домашних хозяйств и других обследований. Такие проблемы включают в другие обследования по двум причинам. Во-первых, цель может заключаться в получении "объясняющих" переменных, полезных для составления комбинационных таблиц и анализа характеристик, которые служат основой для проведения обследования. Хотя в этом случае обследование не ориентировано на получение показателей рабочей силы как таковых, оно может быть полезным для лучшего понимания взаимосвязей между рабочей силой и другими характеристиками, например, рождаемостью, уходом за детьми, здоровьем, доходом, структурой потребления и т.п. И действительно, некоторые вопросы так тесно связаны с характеристиками рабочей силы, что их изучение может потребовать обращения также к этим характеристикам. Например, изучение миграции может потребовать включения таких вопросов, как статус в области активности и занятости, занятие, отрасль, сектор экономики и доход, как до, так и после миграции, с тем чтобы лучше понять ее причины и следствия.

Во-вторых, и это существеннее в настоящем контексте, цель включения "модуля" рабочей силы в другие обследования может заключаться в том, что данные о рабочей силе представляют интерес сами по себе. Включение основных данных в другие проводимые обследования может быть весьма экономичным способом получения существенной информации о рабочей силе и ее характеристиках. Это представляется особенно целесообразным, когда полномасштабное обследование рабочей силы невозможно из-за ограниченных ресурсов или наличия более важных работ. Модуль с вопросами о рабочей силе можно включать в другие обследования от случая к случаю, как это, например, было сделано в рамках Национального обследования семейных бюджетов в Гватемале в 1979—1981 гг. Или это можно сделать на более регулярной основе, как, например, в национальных социально-экономических обследованиях в Индонезии. В ряде стран создана система обследований, которая обеспечи-

ваает достаточно полный и регулярный охват основной проблематики рабочей силы.

Следует, однако, отметить и некоторые недостатки такого подхода. Во-первых, существуют пределы для числа и детализации вопросов о рабочей силе, которые можно включить в другие обследования, связанные преимущественно с другими проблемами. При переписи населения, например, каждый из немногих пунктов о рабочей силе, которые можно включить в анкету, следует свести к одному простому вопросу. Необходимо позаботиться о том, чтобы такие добавления не повлияли отрицательно на общее качество информации, полученной в ходе переписи или обследования, в результате чрезмерной нагрузки на респондентов, задержки в обработке данных и других последствий увеличения объема и сложности работы.

Во-вторых, чтобы обеспечить качество данных и пользу от результатов, необходимо, чтобы различные темы, включенные в одно обследование, были сопоставимы в отношении понятий, определений, методов, отчетных периодов, охвата и других аспектов. Добиться такой сопоставимости не всегда возможно или легко (см., например, анализ проблемы соотнесения объема занятости и экономического благосостояния в гл. 8). В то же время само требование сопоставимости с другими темами, охваченными в том же обследовании, способно отчасти обесценить полученные данные о рабочей силе. Кроме того, обследования могут относиться лишь к некоторым группам населения, поэтому данные о рабочей силе будут лишь служить объясняющими переменными для других явлений. Широко известным примером такой ситуации являются данные о стаже работы супружеских пар, полученные в рамках Обследования рождаемости в мире (1975 г.); какова бы ни была ценность этих данных в качестве объясняющих переменных, они дали мало непосредственной информации о характеристиках рабочей силы для всего населения.

Обследования рабочей силы как средство получения других данных

Организация непрерывного обследования рабочей силы может представлять собой крупное и относительно дорогостоящее начинание. Но когда оно налажено, обследование можно эффективно использовать для охвата дополнительных проблем и поддержки различных обследований домашних хозяйств в других областях. Примером служит Обследование рабочей силы Канады (Canada, 1976). Обследование первоначально было организовано в 1945 г. для получения квартальных оценок характеристик рабочей силы на обще-

национальном уровне. В 1952 г. оно стало ежемесячным обследованием, которое обеспечивало некоторые оценки на субнациональном уровне (в разрезе территорий). Постепенно большинство проводимых в Канаде обследований домашних хозяйств стали использовать возможности обследования рабочей силы и в различной степени интегрироваться с ним по своей структуре и организации. Сюда относятся обследования, проводимые в качестве дополнения к Обследованию рабочей силы, например, повторяющиеся обследования оснащения домашних хозяйств, финансов, потребителей, квартплаты; а также обследования, обеспечивающие более детальную информацию о населении и рабочей силе — об организации труда, наличии рабочих мест, добровольных общественных работах, невыходах на работу, финансах студентов, образовании, путешествиях, привычках курения и проведении досуга. Кроме того, в более независимых обследованиях использовали те же выборочные районы, что и в рамках обследования рабочей силы, но с другими совокупностями домашних хозяйств и иными периодами обследования. Примером этого рода являются обследования семейных бюджетов и семейных расходов. Другие обследования типа обследования здравоохранения в городах сэкономили расходы на организацию выборки, прибегнув к спискам Обследования рабочей силы, хотя они использовали независимую выборку районов и домашних хозяйств. Другими широко известными примерами служат текущее обследование населения в США, которое, помимо рабочей силы, охватывает значительный круг разнообразных вопросов (United States, 1978), и обследование рабочей силы в Австралии (Australia, 1985).

Широкие многоцелевые обследования

Интеграция может принимать также более законченные формы, когда значительное число тем объединяют в одном многоцелевом начинании. В принципе главное преимущество таких всеохватывающих многоцелевых обследований заключается в возможной экономии, связанной с масштабами мероприятия, а также взаимной проверке и совместном анализе детальных данных по различным проблемам. При условии достаточных объемов выборки многоцелевые обследования могут дать массу информации о характеристиках рабочей силы в сочетании с другими аспектами. Примерами являются Социально-экономическое обследование в Индонезии и Национальное исследование семейных бюджетов в Бразилии, проведенное в 1974 г. Во втором случае была собрана информация на основе выборки из 55 тыс. семей (это была подвыборка проводивше-

гося тогда квартального обследования рабочей силы), относившаяся к широкому кругу социально-экономических характеристик, участию в рабочей силе, семейным бюджетам, расходам, питанию и т.д. (Brazil, 1980). В качестве другого примера назовем многоцелевые многораундные обследования дохода, потребления, занятости и ряда других аспектов, которые были недавно организованы в нескольких странах южнее Сахары в связи с проектом Всемирного банка "Социальные аспекты преобразований".

Следует, однако, отметить, что комплексным многоцелевым обследованиям могут быть, а иногда и на деле присущи серьезные недостатки, особенно в более сложных условиях развивающихся стран. Как указывается в "Справочнике по обследованиям домашних хозяйств" ООН (United Nations, 1984, paras. 1.20 — 1.28), к этим недостаткам относятся: продолжительность и сложность опроса; возросшая нагрузка на респондентов; возможное увеличение числа ошибок, не связанных с выборкой; менее эффективное построение анкеты применительно к каждой теме опроса из-за необходимости компромисса между различными требованиями; опасность задержки и сбоев на этапе обработки данных в связи с возросшими объемами и сложностью полученных данных. Остроту некоторых из перечисленных проблем можно ослабить, например, путем использования для различных тем разных подвыборок в дополнении к некоторым ключевым вопросам, относящимся ко всей выборке (в этом случае, разумеется, не все вопросы можно увязать на микроуровне), или организации обработки данных отдельно по различным темам. Во всяком случае требуется осторожность во избежание чрезмерного усложнения или перегрузки обследования.

Система обследований различной ориентации

Еще одна схема интеграции, которая может особенно подходить для условий и требований ряда развивающихся стран, заключается в организации продолжающейся системы обследований с различным содержанием раундов. Каждый раунд может охватывать определенный период, например год, и отдельную предварительную выборку. Система обследований может опираться на общие организационные рамки, персонал и технические средства, однако содержание работы от раунда к раунду меняется, если не считать некоторых ключевых вопросов, которые могут оставаться неизменными. Такая система может включать периодические всесторонние обследования экономически активного населения в качестве основного ориентира раундов. Всестороннее обследование способно дать

структурную или подробную информацию, представляющую долговременный интерес. В приведенном варианте заложено много преимуществ с точки зрения работы по обследованиям в развивающихся странах.

Иллюстрацией такой системы обследований может служить Национальное выборочное обследование в Индии. Непрерывное обследование состоит из годовых раундов, каждый из которых ориентирован на ряд определенных, но взаимосвязанных тем. Например, в годовых раундах примерно каждые пять лет охвачены такие темы, как занятость, безработица, сельская рабочая сила и смежные проблемы. Хотя такая система не обеспечивает постоянного притока текущих данных или информации об изменениях в течение коротких отрезков времени, она хорошо подходит для получения периодической информации о важных структурных характеристиках, а также о долговременных сдвигах. Важными преимуществами описанной системы является то, что она: а) способна сбалансированным образом охватить многообразные взаимосвязанные проблемы, вместо того чтобы расходовать ограниченные ресурсы на изучение одной проблемы в ущерб другим; б) позволяет сосредоточиться на структурных характеристиках и долговременных сдвигах, удовлетворяя тем самым потребности страны в соответствующих данных. Названные преимущества делают этот вариант привлекательным для многих развивающихся стран. Примеры аналогичного подхода включают непрерывную интегрированную программу обследований домашних хозяйств в Ботсване, Национальную программу выборочного обследования и оценки в Кении, а также другие программы обследований в Африке, Азии и Латинской Америке, организованные в сотрудничестве с Национальной программой обследований домашних хозяйств ООН.

Общие черты обследований экономически активного населения

Практика проведения обследований рабочей силы или, в более общем виде, обследований экономически активного населения в странах весьма различна в зависимости от конкретных требований к данным, а также условий и возможностей организации обследований. Однако на основе имеющейся эмпирической информации¹ можно выявить ряд присущим им общих черт.

¹ Помимо обзора национальной практики МБТ (ILO, 1986) имеется анализ регионального опыта проведения обследований рабочей силы ООН (United Nations, 1980, part III); более подробное описание для стран Азии и бассейна Тихого океана содержится у Рао (Rao, 1985).

1. Обследования рабочей силы, как правило, представляют собой крупномасштабные обследования всего населения; часто они являются общенациональными по охвату и имеют тот или иной официальный статус. Такие крупные мероприятия, особенно в развивающихся странах, могут, как правило, проводить только национальные статистические управления или другие крупные государственные органы, ведающие статистической работой. Во многих случаях главными потребителями информации также являются основные государственные органы. Например, из 44 стран, по которым приведена соответствующая информация в исследовании МБТ (1986 г.), в 39 случаях организация обследований целиком или частично лежала на национальных статистических управлениях; практически во всех странах обследования осуществляли статистические подразделения или управления министерства труда. Аналогичным образом, 38 обследований были общенациональными по охвату, если не считать второстепенные исключения; в остальных случаях они относились только к городскому сектору¹.

2. Во многих развитых, а также развивающихся странах обследования рабочей силы предпринимают на непрерывной основе с целью получения информации о сложившихся уровнях и изменениях. Упомянутое выше исследование МБТ имело своей целью охватить все страны, где обследования рабочей силы проводят ежегодно или чаще. В развитых странах преобладали месячные и квартальные обследования. В значительном большинстве развивающихся стран, включенных в исследование, сообщили, что они проводят обследования рабочей силы ежеквартально или реже. Страны, которые осуществляют всесторонние обследования экономически активного населения на случайной, а не постоянной основе, в работе МБТ не учтены. В ряде стран проводят обследования с целью получить более подробную структурную информацию, представляющую долговременный интерес, так как они дают важные данные для планирования и формулирования политики, особенно в развивающихся странах. Видимо, достоинства этой информации не всегда оценивались в полной мере, и в ряде случаев слишком много усилий прилагается для получения текущих данных в отличие от структурной информации долговременного характера.

3. В большинстве стран информацию, относящуюся к домашним хозяйствам и индивидам, в ходе обследований рабочей силы получа-

¹ Как и в любом исследовании, основанном на неполных ответах, в результатах, сообщенных МБТ, возможно смещение. Не исключено, что в странах, не приславших ответ, было больше случаев меньшего охвата обследованием населения страны.

ют путем устного опроса работниками, посещающими респондентов в частных домашних хозяйствах. Другие методы сбора данных, например, путем телефонного опроса, в развивающихся странах во многом непригодны. Даже в развитых странах обследования рабочей силы по почте столкнулись с серьезными трудностями в виде неполучения ответов и низкого качества данных, однако телефонные опросы получают все более широкое распространение.

Охарактеризованные выше общие черты обследований рабочей силы имеют ряд последствий с точки зрения их организации и проведения. Во-первых, из-за официального статуса и общенациональных масштабов обследований рабочей силы к ним, вероятнее всего, предъявляют более жесткие требования в отношении сроков, точности и непротиворечивости данных, особенно сопоставимости динамических рядов, полученных в ходе непрерывных обследований. Указанные требования можно выполнить лишь при вероятностных выборках достаточно большого объема, полученных на основе исходной базы, которая охватывает все население и представлена не только в географическом отношении, но и с учетом сезонных и иных колебаний во времени. Оценки на базе обследований должны быть максимально совместимы с данными из других официальных источников. Многие пользователи требуют агрегированных данных по населению в целом (в отличие от оценок средних величин, долей и коэффициентов и в дополнение к ним), а также оценок динамики этих агрегированных величин: А это, как правило, предполагает использование контрольных итоговых цифр, полученных вне рамок обследования.

Во-вторых, во многих случаях орган, осуществляющий обследование рабочей силы, занят также другими обследованиями и статистическими работами. Это увеличивает важность координации и интеграции при планировании и проведении обследований.

В-третьих, в связи с необходимостью личного опроса время и расходы, связанные с переездами интервьюеров, часто составляют основную статью общих расходов по обследованию. Поэтому выборку следует сделать компактной с помощью процедуры многоступенчатого отбора с учетом географического фактора. Многие операции приходится децентрализовать, что увеличивает необходимость в более тщательном обучении интервьюеров, более строгом контроле качества и т.д.

Но в то же время из-за размеров, регулярности и повторяющегося характера работы существует опасность рутинности и отказа от экспериментирования и новаторства. Поэтому необходимо уделять должное внимание систематической оценке и периодическому пересмот-

ру методов и процедур обследования. Это далеко не простая задача. Проблема может усложниться, если, как это часто случается, имеющиеся ресурсы ограничены и существует нажим с целью увеличить объем (разнообразие, количество) данных, собираемых на основе рутинных методов, в ущерб оценке и повышению их качества. Жесткие графики сбора данных способны еще больше усугубить проблему контроля качества.

4. ОСНОВНЫЕ ПРОБЛЕМЫ ПЛАНИРОВАНИЯ ОБСЛЕДОВАНИЙ

Выбор структуры и масштабов обследования

Измерение экономической активности — сложная задача. Сказанное особенно справедливо в случаях, как это происходит во многих развивающихся странах, где экономическая активность по своему характеру нерегулярна, неорганизована, случайна и не очень-то отличима от неэкономической деятельности. Поэтому на практике любое обследование экономически активного населения должно преследовать вполне конкретные цели и отвечать четко сформулированным требованиям в отношении данных. Этот очевидный момент следует подчеркнуть, так как в менее развитых странах иногда прибегают к некритически импортированным и заимствованным методам обследования из статистически более развитых стран, где они были разработаны в других условиях и часто применительно к другим потребностям в информации.

Структура обследования

Как отмечено в предыдущем разделе, выбор типа обследования в основном следует сделать между: а) непрерывной операцией, направленной на получение динамических рядов данных о текущем уровне занятости, безработицы и неполной занятости и информации о краткосрочных изменениях, и б) случайными или менее частыми, но, возможно, более детальными исследованиями для получения информации более структурного свойства и характеризующей сдвиги за более продолжительное время. Чтобы данные отвечали потребностям, указанным в международных стандартах (они охарактеризованы в начале предыдущего параграфа), могут понадобиться оба типа обследований. Однако страна с ограниченными ресурсами может иногда быть вынуждена ограничиться одним из них или во всяком случае уделять одной системе больше внимания, чем другим, в зависимости от конкретной ситуации и требований к данным.

С этим же связана проблема выбора измерителя (или измерителей) экономически активного населения: это может быть “населе-

ние, активное в данный момент”, применительно к краткому отчетному периоду в одну неделю или один день (см. гл. 3) или “обычно активное население” применительно к продолжительному отчетному периоду, например году (см. гл. 4), либо сочетание этих показателей. Например, когда требуется динамический ряд данных о текущих уровнях и краткосрочных изменениях, больше подходит непрерывный тип обследования с кратким отчетным периодом. Но если полезнее (как это бывает во многих развивающихся странах, особенно в сельской местности) иметь более детальную информацию об обычно активном населении и его характеристиках, а также о долговременных сдвигах в нем, тогда больше подойдут менее частые, но более подробные обследования на основе продолжительного отчетного периода. Разумеется, когда это возможно, следует прибегнуть к сочетанию различных моделей с целью получить более обширные данные, полнее удовлетворяющие различные потребности в информации. В частности, независимо от того, выбрана ли непрерывная или случайная модель, может оказаться целесообразным сочетать измерение обычно активного населения и населения, активного в данный момент, с тем чтобы получить комбинационные таблицы, характеризующие статус индивидов.

Какова бы ни была структура обследования, важно учесть требования в отношении оперативности информации. При непрерывном типе обследования важными аспектами являются: частота сбора данных; частота опубликования оценок (которая может отличаться от частоты сбора данных); относительная важность различных видов оценок (например, оценок текущего уровня, тенденций или чистых изменений от одного периода к другому, общие изменения отдельных показателей и т.д.). Аналогичным образом менее частые и более подробные обследования структурного характера могут потребовать длительного времени или во всяком случае правильного учета сезонных и иных колебаний, происходящих с течением времени. Кроме того, для правильного определения времени проведения работы приходится учитывать особые потребности и условия, связанные, например, с разработкой плана развития или проведением переписи населения.

Масштабы обследования

Определение масштабов обследования относится к числу принципиальных решений в процессе его планирования. Оно зависит почти от всех прочих аспектов и в свою очередь влияет на них, включая способ получения данных, организацию операции и обработки дан-

ных, требования к количеству и квалификации персонала, время и точность выдаваемой информации и, прежде всего, стоимость обследования. Следует отметить, что под “масштабами” понимают не только объем выборки; скорее речь идет о характере совокупных усилий, требующихся для проведения обследования. Они зависят от ряда факторов, например:

- число обследуемых единиц (объем выборки);

- доступность единиц, что определяется, в частности, распределением и характеристиками населения, наличием средств связи, методом получения данных, типом выборки и др.;

- сложность и объем информации, подлежащей сбору, а также, что часто еще важнее, требующаяся обработка;

- частота сбора данных и выдачи результатов обследования;

- время, отведенное для завершения работы, что определяет темпы ее проведения.

При планировании обследования необходимо должным образом учитывать масштабы работы. Очевидно, что при любом обследовании объем выборки, а также количество и частоту сбора данных необходимо определить в зависимости от содержательных целей обследования. Объем выборки должен быть достаточно большим, чтобы ошибки выборки были в некоторых пределах, что позволит получить в результате обследования информацию, полезную для целей исследования и принятия решений. Но точность выборки — это лишь одна составляющая качества обследования. Результаты подвержены не только ошибкам выборки, но и ошибкам, связанным с охватом, содержанием и проведением работы, и обусловленным различными причинами. Это ошибки регистрации, они зависят от качества организации работы, а это в свою очередь — отчасти от ее масштабов. Важно обеспечить, чтобы объем выборки находился в пределах управляемости. Попытки охватить чрезмерно большие выборки обычно приводили к возникновению проблем во многих обследованиях.

Объем и сложность собираемой информации следует определять прежде всего с учетом требований пользователей. Они могут требовать комплексных многотемных обследований. Используя существующее обследование для получения дополнительной информации, можно добиться экономии, обусловленной масштабами работы. Точно так же информацию по комплексным проблемам типа экономической активности, к тому же в сложных и разнообразных условиях развивающихся стран, можно качественно получить лишь на основе тщательно проведенного опроса и хорошо сформулированной анкеты. Но сверх некоторого предела чрезмерная нагрузка

на счетчиков и респондентов может отрицательно отразиться на качестве данных.

Взаимосвязь между объемом выборки и стоимостью обследования осознана хорошо, чего нельзя сказать о взаимосвязи между объемом выборки и оперативностью данных. Оперативность важна при любом обследовании, но она может оказаться решающим параметром, когда обследование носит постоянный характер, а государственные аспекты мероприятия требуют жесткого графика выдачи результатов. Масштабы, сложность и частоту обследования необходимо ограничить, с тем чтобы сбор данных происходил темпами, при которых их можно своевременно обработать. В противном случае при непрерывной работе будут все больше накапливаться необработанные материалы, что обесценит сбор текущей информации. Не менее важное соображение при организации непрерывного обследования заключается в следующем: его масштабы необходимо определить таким образом, чтобы обеспечить устойчивое продолжение динамического ряда данных, для которого оно было задумано. Сказанное предполагает наличие некоторых избыточных мощностей на случай непредвиденной нагрузки или неблагоприятных обстоятельств.

Планирование и организация обследования

Хотя обследованиям экономически активного населения присущи специфические особенности и требования, к ним также применимы основные принципы проектирования и организации крупномасштабных обследований домашних хозяйств. В части I "Справочника по обследованиям домашних хозяйств" ООН (United Nations, 1984) содержится подробное описание общих принципов планирования и организации обследования. Данное описание исходит из широкой перспективы системы (или программы) взаимосвязанных обследований, а не из одного изолированного обследования. Это делает документ еще более ценным с точки зрения обследований экономически активного населения, которые, как отмечалось выше, обычно необходимо планировать и проводить в тесной взаимосвязи с другими обследованиями. Более конкретно вопросы планирования программ интегрированных обследований домашних хозяйств рассмотрены в работе ООН (United Nations, 1986). Опираясь на указанные источники, ниже обобщены некоторые важные моменты планирования и организации обследования.

Обследования домашних хозяйств — это существенная часть национальной системы статистики, но это не вся система. Поэтому планирование обследований домашних хозяйств следует осуществ-

лать в рамках общих статистических планов и приоритетов. Взаимосвязь между обследованиями домашних хозяйств и другими источниками данных необходимо уяснить на ранней стадии процесса планирования, как только определены требования к данным. Следует оценить различные источники данных с точки зрения их способности удовлетворить потребность в информации, связанных с ними издержек, требований к технике и квалификации персонала при различных вариантах. Как правило, более целесообразно найти сочетание взаимодополняющих источников, используя преимущества каждого из них, нежели остановиться только на одном источнике данных.

Особое значение имеет связь между обследованиями домашних хозяйств и переписями населения и жилья. При разработке любого плана обследования необходимо должным образом учитывать потребность в обучении, особые условия и объем работы при переписях. Хотя в период проведения переписи обследования, возможно, отчасти придется свернуть (особенно если оба мероприятия осуществляет один орган), завершение переписи способно дать большой толчок обследованиям. Период, непосредственно следующий за переписью, — это наиболее удобный момент для тщательного переосмысления и серьезной перестройки существующей системы обследования или для начала создания такой системы.

Планирование обследования включает следующие конкретные шаги:

- выявление подлежащих охвату проблем, их сравнительной значимости и срочности;

- выявление особых требований к структуре;

- группировка проблем с учетом их содержательных и организационных особенностей и определение на этой основе отдельных обследований или их раундов;

- определение времени, частоты, масштабов, охвата населения и общей структуры каждого обследования;

- разработка плана мероприятий при должном учете условий и ограничений, относящихся к таким факторам, как характеристика обследуемого населения, доступ к элементам выборки и наличие различных ресурсов, включая материалы выборки, персонал на разных уровнях, мощности для обработки данных (United Nations, 1986, p. 20 — 38).

Относительно организации обследования в “Справочнике по обследованиям домашних хозяйств” ООН (United Nations, 1984, paras. 2.2 — 2.14) отмечается, что независимо от характера организации национальной статистической службы для проведения лю-

бого обследования требуются некие минимальные административные и технические предпосылки. Любое обследование предполагает наличие по меньшей мере какого-то ядра специалистов, способных осуществить ряд работ, включая:

- плановые и управленческие функции, в том числе определение приоритетов, разработку графиков, распределение ресурсов и персонала и т.п.;

- конкретизацию проблематики обследования, в том числе деталей содержания, понятий и определений, разработку анкеты, выходных таблиц и планов анализа;

- определение выборки и смежных аспектов структуры обследования;

- организацию сбора данных, в том числе пробных проверок, подбор и обучение персонала, наблюдение за его работой, контроль качества;

- обработку данных, в том числе системное проектирование, программирование, редактирование, кодирование, табулирование и хранение данных;

- обслуживающие функции, например, работу компьютеров, печатание и выдачу документов.

Техническое проектирование

Техническое проектирование обследования — это процесс, предполагающий одновременный учет различных соображений и требований, а также компромисс между ними. Оно начинается с определения содержательных целей, включая характер требуемой информации, частоту ее поступления, обследуемое население, группы населения, по которым необходимы отдельные оценки, требования к оперативности и точности, приоритеты среди целей. Это подскажет возможные способы сбора информации, общую структуру и масштабы операции и возможные издержки ее проведения. Издержки следует оценивать в свете практических требований к проведению обследования при преобладающих условиях: имеющихся финансовых, технических и кадровых возможностях; временных факторах; а также, особенно при крупномасштабном мероприятии, с учетом возможностей увязки с другими операциями, как в организационном, так и содержательном смысле. В течение всего процесса могут потребоваться корректировки с целью примирить теоретически желаемое с практически возможным.

Различные аспекты проектирования обследования подробно рассмотрены в последующих главах. Они включают уточнение структуры обследования и определение выборки (гл. 11); разработку ан-

кеты (гл. 12); установление процедур и методов предварительной проверки, сбора и обработки данных (гл. 13); оценку и контроль качества данных (гл. 14). Важным элементом процесса является пересмотр и корректировка инструментов и процедур обследования, которые периодически проводят на основе накопленного опыта и новой информации и требований; эти вопросы рассмотрены ниже.

5. Реорганизация обследований

Во многих странах обследования рабочей силы уже проводят, и периодически их следует совершенствовать. Очевидно, что требования совершенствования прежде всего относятся к обследованиям непрерывного типа, однако многие рассмотренные моменты применимы также к проводимым менее часто случайным обследованиям, так как и их организация со временем нуждается в улучшении. Общие принципы проектирования обследований относятся также к их реорганизации. Однако при совершенствовании существующей системы необходимо учитывать ряд дополнительных факторов. Дело в том, что периодическая реорганизация, важная с точки зрения улучшения и даже сохранения системы и ее эффективности, может иметь отрицательные последствия. Например, введение новых понятий и процедур способно повлиять на преемственность и сопоставимость имеющихся динамических рядов. Аналогичным образом реорганизация может потребовать новых крупных затрат на подготовку материалов, обучение персонала, разработку программ для обработки данных и т.п. Поэтому необходим компромисс между потребностью в постоянном улучшении организации и процедур обследования, с одной стороны, и необходимостью сохранить сопоставимость и ценность имеющихся мощностей, с другой.

Цели реорганизации

Существующая организация и процедуры обследования могут нуждаться в модификации по ряду причин.

Во-первых, могут возникнуть изменения в содержательных целях обследования. Эти изменения могут относиться к масштабам и охвату; определениям и используемым понятиям; изучаемым проблемам; характеру, частоте и точности требуемых оценок; изучаемым группам населения; географическим и иным группировкам результатов.

Во-вторых, реорганизация обследования может оказаться необходимой из-за изменений в имеющихся ресурсах, условиях работы и

других внешних факторах. Когда эти факторы удается предусмотреть, можно запланировать и изменение процедур обследования. Но это не всегда так. Нередко, к примеру, статистический орган сталкивается с резким сокращением своего бюджета, причем происходит это внезапным и непредсказуемым образом. Иногда появляются дополнительные ресурсы, так как в другом месте были неожиданно сэкономлены средства. Серьезные коррективы в проведении обследования могут быть обусловлены засухой или другими климатическими факторами, соображениями безопасности и т.п. При организации непрерывных обследований в развивающихся странах часто недостаточно учитывали возможные негативные последствия такого рода явлений. Эти последствия можно уменьшить путем принятия гибких проектов и процедур, позволяющих быстро приспособиться к изменившимся условиям и требованиям.

В-третьих, применительно к крупному продолжающемуся мероприятию типа непрерывного обследования рабочей силы весьма важным соображением является возможность его использования как средства сбора информации по другим проблемам и для других обследований. В ряде случаев в результате ряда реорганизаций обследования рабочей силы постепенно превратились из более или менее специальных мероприятий в обследования общего многоцелевого характера.

В-четвертых, и это наиболее частая причина, цель реорганизации обследования состоит в улучшении качества данных и снижении затрат, что достигается с помощью более эффективной организации и совершенных процедур, основанных на более свежей информации о размере и характеристиках изучаемого населения, на более качественной выборке, на анализе ошибок выборки и регистрации, а также издержек и проведения обследования. Самый подходящий случай для серьезной реорганизации такого рода представляется при наличии результатов новой переписи населения. Результаты переписи могут позволить обновить выборку с учетом роста населения и изменения его характеристик, а также коррективов в границах переписи и других статистических или административных моментов. Объем выборки и ее распределение можно скорректировать в свете изменений и географии обследования. Усовершенствования в сборе и обработке данных и методах оценки возможны также на основе накопленного опыта, а также улучшений технологии и имеющихся технических средств.

Отдельно заслуживает упоминания то, что иногда возникновение (или обнаружение) существенных неувязок в действующей системе может вызвать необходимость коренного пересмотра охвата, содер-

жания, масштабов, структуры и процедур обследования или даже всего способа сбора данных. Такие изменения могут означать организацию более или менее нового обследования, а не просто улучшение существующего.

Стратегия реорганизации

Как отмечалось выше, реорганизация предполагает компромисс между необходимостью усовершенствования обследования и потребностью обеспечить преемственность и минимизировать отрицательные последствия изменения сложившихся процедур. Преемственность может быть необходима для сохранения сопоставимости динамических рядов, предупреждения административных или организационных неувязок (их последствия могут усугубляться, если время и персонал для реорганизации ограничены) и поддержания ценности уже осуществленных инвестиций.

Указанные соображения не означают, что перемен следует вообще избегать или сводить их к минимуму. Более или менее серьезная реорганизация, конечно, периодически важна в любом продолжающемся обследовании, так как невозможно бесконечно сохранять сложившуюся организацию и одни и те же методы. Но в то же время необходимо проявлять осторожность, сводить к минимуму отрицательные последствия перемен и избегать постоянных реорганизаций. В большинстве случаев правильная стратегия заключается в накоплении перемен и осуществлении серьезных реорганизаций лишь периодически, с охватом как можно большего числа аспектов. К тому же перемены следует осуществлять, когда они явно назрели и необходимы.

Хотя нередко желательно сохранить как можно больше элементов существующего проекта и сложившихся процедур, наступают моменты, когда наиболее эффективной стратегией является радикальный, а не постепенный подход к реорганизации.

Отрицательные последствия перемен можно иногда уменьшить, проводя параллельно и одновременно обследования на основе старого и новых проектов и обеспечивая альтернативные оценки и группировки в течение переходного периода. Однако издержки и нагрузка, связанные с такой двойной системой, способны перевесить полученные выгоды, что делает такой подход нецелесообразным. Кроме того, некоторым пользователям не нравится получать две совокупности оценок.

Полезно провести различие между двумя видами реорганизации, хотя возможны и промежуточные или смешанные варианты:

а) Реорганизация, направленная преимущественно на повышение эффективности обследований в смысле представительности, точности выборки, затрат, процедур выполнения работ, качества данных и т.п., без серьезного изменения понятий, определений, содержания, отчетного периода или существенных условий и метода сбора данных. Примерами могут служить: постепенное увеличение объема выборки и даже обследуемого населения; перераспределение выборки между различными районами; более правильное определение выборки и процедур оценки на основе более свежей исходной информации; более эффективная организация сбора данных; улучшение контроля; усовершенствование процедур для ускорения обработки данных.

б) Реорганизация, предпринятая в связи с существенными изменениями целей или организационных условий проведения обследования. Это может быть связано с изменением понятий, определений, отчетного периода, периодичности, характера требуемых оценок, районов, для которых даются оценки, масштабов работы, методов сбора данных и т.п.

Первый вид реорганизации может и не иметь серьезных последствий для преемственности и сопоставимости данных во времени; и ее обычно проводят во многих странах вслед за каждой переписью населения. Наиболее хорошо описанными примерами периодической реорганизации после переписи населения могут, пожалуй, служить текущие обследования населения в США и обследования рабочей силы в Канаде. В этих случаях каждой реорганизации предшествует серьезное изучение для оценки действующего проекта и процедур. Помимо корректировки выборки в связи с демографическими изменениями, детально изучаются многие важные и второстепенные аспекты обследования с целью выяснить возможные области улучшения. К примеру, после переписи 1980 г. текущее обследование населения было реорганизовано с тем, чтобы получать более точные оценки по штатам в дополнение к общенациональным оценкам. Аналогичным образом после переписи 1981 г. в Канаде обследование рабочей силы было пересмотрено не только с целью улучшения выборки (без изменения принципиальных моментов, вроде метода сбора данных и периодичности), но и явно для того, чтобы добиться большей гибкости при использовании обследования рабочей силы в качестве основы для проведения других обследований домашних хозяйств. Большинство обследований домашних хозяйств, проводившихся Статистическим бюро Канады, уже опирались на технические средства обследования рабочей силы. Цель реорганизации заключалась в изучении возможности усилить

имеющиеся связи. (Детальное описание методологии обоих обследований содержится в: United States, 1978; Canada, 1976).

Второй тип реорганизации обычно затрагивает сопоставимость данных в значительно большей степени. Это особенно относится к обследованиям рабочей силы, при которых результирующие оценки, например, уровни участия в рабочей силе, занятости и безработицы, могут очень сильно зависеть от отчетного периода и используемых понятий и определений. Следовательно, изменения в эти аспекты организации и процедур обследования следует вносить с большей осторожностью, а их последствия рассматривать более тщательно.

На примере обследования рабочей силы в Финляндии можно проиллюстрировать коренную реорганизацию всего метода сбора данных. До 1983 г. обследования проводились на основе почтового опроса. Это приводило к неудовлетворительному качеству данных и большой доле неполученных ответов (свыше 30 %). Новым проектом был предусмотрен устный опрос — 90 % по телефону, остальных — лично. Указанные изменения обеспечили улучшение содержания обследования, повышение доли полученных ответов и качества данных (Salmi and Kiiski, 1984).

Обследование рабочей силы в Таиланде может служить примером радикальных изменений понятий и определений, которые, естественно, повлияли также на сопоставимость данных во времени. Обследование было организовано в 1963 г., а с 1971 г. были предусмотрены два раунда в год, для того чтобы охватить отдельно сельскохозяйственный сезон и межсезонье. С 1977 по 1982 гг. при обследовании применили концепцию “использования рабочей силы”, от которой впоследствии по различным причинам отказались. Другие важные изменения в определениях, введенные при реорганизации, затрагивали неоплачиваемых работающих членов семьи; лиц, готовых приступить к работе; лиц, не работающих, но ожидающих начала сезона сельскохозяйственных работ. Ранее неоплачиваемых членов семьи, которые в течение отчетной недели работали менее 20 часов и не хотели работать больше, считали вне рабочей силы; после пересмотра (и с учетом новых международных стандартов) неоплачиваемых членов семьи стали включать в рабочую силу, если они проработали хотя бы один час. Лиц, которые не работали, но утверждали, что готовы приступить к работе, отнесли к “безработным”. Серьезное изменение, затронувшее классификацию более 5 млн. человек, относилось к тем, кто утверждал, что они не были готовы приступить к работе в отчетном периоде, но обычно работают в сельскохозяйственный сезон и ожидают его. В пересмотренном

варианте обследования их включили в совокупную рабочую силу в качестве отдельной категории “сезонно неактивных” (Thailand, 1985).

Следует отметить, что не все формальные изменения понятий и определений обязательно отражаются на сопоставимости динамических рядов. Во-первых, некоторые изменения могут лишь означать формальное признание некоторых методов и опыта, которые уже вошли в практику. Это может произойти, например, когда прежние инструкции и процедуры были нереалистичны и поэтому интервьюеры их единообразно не выполняли. Так, если отчетный период для выявления активно ищущих работу в обследовании ранее был определен неоправданно коротким (например, в одну неделю), вполне возможно, что фактически полученная информация относилась к более реалистичному и продолжительному периоду, даже если он четко не был определен (скажем, несколько недель до опроса). Если такой факт установлен, желательно формально закрепить более продолжительный период, чтобы устранить несоответствие между инструкцией и практикой, а также между подходами различных интервьюеров.

Во-вторых, часто удается восстановить сопоставимость путем выделения изменившихся категорий в отдельные подгруппы, причем некоторые из них могли и не быть затронуты в результате изменений. Например, если определение “безработные” было расширено, с тем чтобы оно включало не только тех, кто “ищет работу и готов к ней приступить”, но и тех, “кто готов приступить к работе, но не ищет ее”, то первую группу можно по-прежнему табулировать отдельно и сравнивать с результатами на основе прежних определений. Аналогичный пример приведен выше по Таиланду применительно к подгруппе “сезонно неактивных”.

Литература

- Australia. 1985. *Labour Force Survey Australia*. Canberra, Australian Bureau of Statistics, Catalogue No. 6203.0.
- Brazil. 1980. *ENDEF: Obrjectivos, Metodologia e Dicionario do Banco de Informaçoes*. Rio de Janeiro, Fundacao Instituto Brasileiro de Geografia e Eststistica.
- Canada. 1976. *Methodology of the Canadian Labour Force Survey*. Ottawa, Statistics Canada, Catalogue No. 71—526 Occasional.

- International Labour Office (ILO). 1983. *Thirteenth International Conference of Labour Statisticians: Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*.
- 1986. *Statistical sources and methods*. Vol. 3, Economically active population, employment, unemployment and hours of work (household surveys). Geneva.
- 1988. *Current international Recommendations on labour statistics*, Geneva.
- Kish, L.; Verma, V. 1986. "Complete censuses and samples", in *Journal of Official Statistics*, 2 (4), pp. 381—395.
- Murthy, M. N.; Roy, A. S. 1970. "A problem of integration of surveys: A case study", in *Journal of the American Statistical Association*, 65, pp. 123—135.
- Rao, M. V. S. 1985. *Labour force and family living surveys in Asia and the Pacific*. Bangkok, Economic and Social Commission for Asia and the Pacific (ESCP), Statistics Division, St/Escap/312.
- Salmi, H.; Kiiski, S. 1984. "Developing the quality of the labour force survey: The Finnish approach", in *Statistisk Tidskrift* (Stockholm, Statistics Sweden), 1984. 1, pp. 16—26.
- Thailand. 1985. *Report on the Labour Force Survey*. Bangkok, National Statistical Office.
- United Nations. 1980. *Principles and Recommendations for population and housing censuses*. Series M, No. 67, New York, UN Statistical Office.
- 1984. *Handbook of household surveys*. Series F, No. 31, New York, UN Statistical Office.
- 1986. *Sampling frames and sample designs for integrated household survey programmes*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.
- United States. 1978. *The current population survey: Design and methodology*. Technical Paper, No. 40. Washington, DC, US Bureau of the Census.
- World Fertility Survey. 1975. *Core Questionnaires*, WFS Basic Documentation, No. 1. The Hague, International Statistical Institute.

Глава 11. ПЛАНИРОВАНИЕ ВЫБОРКИ

1. ВВЕДЕНИЕ

В настоящей главе рассмотрены технические и практические вопросы планирования выборки и оценки показателей, представляющие особый интерес применительно к обследованиям рабочей силы.

Планирование выборки для любого обследования зависит от конкретных требований к данным и условий его проведения. В гл. 10 описаны общие черты и схемы обследований рабочей силы. Эти общие черты, применимые ко многим обследованиям, несмотря на существенные различия в практике отдельных стран, следует иметь в виду при обсуждении планирования и организации выборки.

Необходимо подчеркнуть, что в настоящей главе рассмотрены лишь некоторые вопросы выборочного метода, относящиеся к обследованиям рабочей силы, причем особое внимание уделено условиям и потребностям развивающихся стран. Здесь не затронуты общие принципы выборочного метода, которые подробно рассмотрены во многих руководствах и учебниках по этому вопросу. В качестве стандартных справочных изданий можно назвать книги Йейтса (Yates, 1949), Хансена, Гурвица и Мэдоу (Hansen, Hurwitz and Madow, 1953), Кокрена (Cochran, 1953), Сикатме (Sukhatme, 1954), Деминга (Deming, 1960), Киша (Kish, 1965) и Мэрфи (Murthy, 1967). Кроме того, следует упомянуть издания ООН: "Краткий справочник по выборочному методу" (1960), "Справочник по обследованиям домашних хозяйств" (1984) и "Основа и планирование выборки для интегрированных программ обследований домашних хозяйств" (1986).

Материал настоящей главы расположен следующим образом.

В парагр. 2 рассмотрены вопросы, относящиеся к выявлению обследуемого населения, на основе которого планируется выборка. Парагр. 3 посвящен требованиям к выборке, с тем чтобы она представляла изучаемое население. Рассмотрено также смежное понятие главной выборки, которую можно использовать для согласования выборок при проведении различных обследований и раундов. Некоторые практические аспекты планирования выборки кратко освещены в парагр. 4. При этом не предпринималась попытка обстоятельно изложить сложную теорию и практику планирования выборки для крупномасштабных обследований домашних хозяйств. Парагр. 5 посвящен временному аспекту выборки, что является существенным моментом при многих обследованиях домашних хозяйств. Процедуры оценки рассмотрены в парагр. 6. Любое непрерывное обследование предполагает периодическую оценку и переоценку целей и методов, а следовательно, периодическое утвержде-

ние и пересмотр процедур. Пересмотр выборки — это один из наиболее важных аспектов этого процесса, и ему посвящен заключительный (7) параграф главы.

2. ОПРЕДЕЛЕНИЕ ОБСЛЕДУЕМОГО НАСЕЛЕНИЯ

Определение населения, на которое следует распространить результаты выборочного обследования — это принципиальный аспект планирования обследования. Подлежащее охвату население было охарактеризовано в гл. 2. Было отмечено, что обследованиями экономически активного населения следует в принципе охватить все население независимо от статуса активности, пола, семейного положения, этнической принадлежности и т.д., однако, на практике могут оказаться необходимыми некоторые ограничения. Например, охват многих обследований приходится ограничить гражданским населением вне специальных учреждений, т.е. исключить военнослужащих и тех, кто находится в некоторых заведениях. По практическим соображениям иногда исключают иностранцев, кочевников, сезонных мигрантов, бездомных, лиц, проживающих в отдаленных и труднодоступных районах, и еще некоторые группы населения. В домашних хозяйствах, охваченных обследованием, как правило, не учитывают детей ниже определенного минимального возраста, так как вопрос об экономической активности применительно к ним не имеет смысла.

В настоящем разделе различные проблемы, связанные с определением обследуемого населения, рассмотрены прежде всего под углом зрения процедур планирования выборки. Принципиальные решения о характере и охвате населения обследованием часто принимают на более ранней стадии планирования работы, но на стадии технического проектирования эти вопросы требуют дальнейшего уточнения. Такие уточнения затрагивают:

особенности населения, т.е. виды и характеристики составляющих его элементов;

расположение населения в пространстве, т.е. границы географического охвата, и во времени, т.е. период, к которому население относится.

Особенности населения

При обследованиях рабочей силы, охватывающих все население, элементами для изучения, как правило, служат отдельные лица и их определенные характеристики, а также домашние хозяйства и другие социальные образования, в которых индивиды проживают. Чтобы определить подлежащее охвату население, необходимо, сле-

довательно, установить: а) тип домашних хозяйств и социальных образований, подлежащих учету; б) правила отнесения индивидов к этим единицам (т.е. определить, что образует "домашнее хозяйство"); в) характеристики, определяющие включение индивидов в охваченное обследование население или исключение из него.

Домашние хозяйства и другие социальные образования индивидов

Наибольший интерес представляют лица, проживающие в *частных домашних хозяйствах*. Определения домашних хозяйств, рекомендуемые в международном масштабе, содержатся в "Принципах и рекомендациях по проведению переписей населения и жилья" (United Nations, 1980). Домашнее хозяйство определяется как "лицо или лица, которые приняли меры, индивидуально или в рамках группы, для обеспечения себя продовольствием и другими предметами первой необходимости". В Рекомендациях проводится различие между домашними хозяйствами из одного лица или нескольких лиц, а последние подразделяются на: базовые (nuclear) домашние хозяйства, каждое из которых состоит из одной семейной единицы; расширенные домашние хозяйства, включающие других лиц помимо базовой семьи; составные домашние хозяйства, включающие лиц, не состоящих в родстве. В Рекомендациях, однако, отмечается, что страны могут признать целесообразным изменить определение и классификацию с учетом национальных особенностей. Так, в ряде стран Африки имеются сельские районы, где люди, как правило, проживают большими группами по 30—40 человек, в этом случае практическое решение состоит в том, чтобы определить домашнее хозяйство как группу лиц, которые "живут и питаются вместе" (United Nations, 1984, para. 20.12). Точно так же в некоторых арабских странах понятие домашнего хозяйства может требовать видоизменения с учетом полигамной ситуации, при которой жены одного мужа могут проживать в отдельных жилищах (ibid., para. 21.18). Во многих других случаях связь между отдельными категориями лиц (жильцы, пансионеры, домашняя прислуга) и домашними хозяйствами может быть не столь очевидна и может требовать особого внимания (ibid., para. 10.64).

Полезно провести различие между ситуациями, когда домашнее хозяйство само по себе является единицей для сбора, табулирования и анализа данных и когда его лишь используют для выявления и отбора лиц, подлежащих учету в ходе обследования. В первом случае выбор конкретного определения может иметь важные и принципиальные последствия. Во втором случае главное требование сводится

лишь к тому, чтобы определение домашнего хозяйства охватывало всех индивидов, представляющих интерес, и чтобы его однозначно толковали во всех раундах обследования и при каждом опросе.

Помимо лиц, проживающих в домашних хозяйствах, необходимо уточнить подход к индивидам, проживающим в различных особых условиях. Практика разных стран неодинакова, и невозможно дать какие-то общие рекомендации, следует лишь подчеркнуть, что необходимо: а) тщательно взвесить основания для исключения особой группы в каждом отдельном случае; б) четко определить исключаемые группы, равно как их размер и соответствующие характеристики; в) попытаться установить последствия таких исключений с точки зрения пригодности и качества результатов обследования; г) максимально возможно обеспечить преемственность охвата во времени и в различных обследованиях, особенно когда обследование рабочей силы составляет часть интегрированной программы обследований. Где это возможно, информацию об исключенных группах, например военнослужащих, необходимо получить из других источников, с тем чтобы дополнить результаты обследования домашних хозяйств.

Правила объединения в группы

Определение обследуемого населения предполагает далее формулирование правил объединения, в соответствии с которыми индивиды с домашними хозяйствами и другими соответствующими образованиями учитываются при организации выборки и анализе данных. Выбор заключается главным образом между определением охвата де-факто и де-юре. При подходе де-юре решающим является постоянное место проживания индивида. Постоянно проживающие лица включаются в обследование, даже если в момент опроса они временно отсутствуют; посетители и гости — исключаются. При подходе де-факто — лица учитываются в зависимости от того, где они находятся в момент опроса, независимо от их постоянного места жительства. Такой подход, как правило, дает большую долю ответов и контактов и лучше всего подходит для краткого периода обследования и в тех случаях, когда обследуемое население весьма мобильно (например, кочевое, бездомные). Подход де-юре обеспечивает более четкую и стабильную картину состава домашних хозяйств. Он более удобен при непрерывном обследовании, особенно когда повторный учет растягивается на значительный период. По указанным причинам он получил большее распространение в обследованиях рабочей силы. Трудности, связанные с подходом де-юре, могут заключаться в выявлении «проживающих обычно» лиц. Име-

ется мало эмпирической информации по вопросу о том, в какой мере выбранный подход отражается на охвате, доле ответов и других аспектах выборочного метода. Анализ результатов обследований в более чем 10 развивающихся странах показал, что в 5 — 10% выборочной совокупности домашних хозяйств в момент обследования находились гости; в то же время было установлено, что, как правило, вдвое больше домашних хозяйств сообщали о временном отсутствии одного или нескольких постоянно проживающих лиц. За этими средними цифрами скрываются значительные колебания по странам (Verma, 1980). Из приведенных результатов вытекают интересные выводы. Во-первых, в то время как временно отсутствующих на основе подхода де-юре выявить ("охватить") можно, их, как правило, непосредственно опросить не удастся, ибо их нет на месте, а следовательно, ответы от них не будут получены. Практика свидетельствует, что количество таких индивидов может быть довольно значительным. Их можно учесть, если допустимо получение информации от других лиц. Во-вторых, подход де-факто обычно используют для того, чтобы уменьшить число тех, кто не охвачен непосредственными контактами. Это особенно важно, когда предусмотрены только прямые контакты, а не ответы других лиц. В принципе подход де-факто позволяет снизить долю неполученных ответов благодаря тому, что временно отсутствующие лица заменяются гостями или посетителями, находящимися в выборочном домашнем хозяйстве. Упомянутый выше анализ показал, что эта "компенсация" весьма относительна, так как вторая группа в среднем наполовину меньше первой. Разность между обеими группами составляют отчасти лица "в движении" ("транзитные"), т.е. пребывающие в момент обследования вне домашних хозяйств; она может также свидетельствовать об ошибочном включении членов семьи, которые фактически проживают в другом месте. "Транзитных" лиц при подходе де-факто не учитывают даже среди тех, кто не был опрошен, что скорее говорит о неполном охвате населения, подлежащего обследованию. Такой неполный охват при обследованиях на основе подхода де-факто может быть довольно значительным и избирательным, хотя нередко это остается незамеченным.

Правила включения

Включение индивидов в обследуемое население обычно обусловлено демографическими и другими индивидуальными характеристиками. Наиболее важной из них является возраст индивида, что предполагает отказ от подробного опроса об экономической актив-

ности детей моложе некоего минимального возраста (а возможно, и пожилых людей старше определенного максимального возраста). Различные соображения, связанные с выбором этих возрастных пределов, рассмотрены в парагр. 2 гл. 2. Было отмечено, что каким бы ни был взят минимальный возраст, в ряде стран может быть значительное количество детей моложе этого возраста, которые заняты той или иной экономической деятельностью. В таких ситуациях может возникнуть необходимость получения дополнительных данных, хотя бы от случая к случаю, о количестве, характеристиках и условиях труда этих детей, имея в виду, например, изучить переходный период от учебы к трудовой деятельности или выявить взаимосвязь между посещением школы и экономической активностью.

Обследование рабочей силы иногда ориентировано на отдельные группы населения, например, трудящихся-мигрантов, иностранных рабочих, учащихся и инвалидов. В этих случаях для правильного определения выборки могут потребоваться особые процедуры отбора для выявления мест концентрации такого рода лиц.

Охват населения

Этот вопрос имеет два аспекта: а) географические границы охвата; б) временные рамки охвата.

Географический аспект

Многие обследования рабочей силы имеют общенациональные масштабы, т.е. охватывают все географические регионы, городскую и сельскую местность страны. Существует, конечно, много случаев, когда необходима работа меньшего масштаба. Во-первых, это может быть обусловлено ограниченными целями обследования. Например, в ряде стран обследования рабочей силы ограничивают только столицей, в других охватывают все крупные города, а в третьих — все города, включая небольшие. Практика исключения из обследований сельской местности распространена в Латинской Америке и странах Карибского бассейна в большей степени, чем в других регионах; примерами могут служить обследования в Аргентине, Бразилии, Колумбии, Мексике, Уругвае и Ямайке (ILO, 1986). Точно так же в некоторых раундах Национального выборочного обследования Индии была охвачена только рабочая сила в городах (без учета ее в сельской местности). Причиной такого подхода могут быть ограниченные ресурсы, но чаще он основан на убеждении, что требования к данным и условия (а следовательно, структура затрат и пригодная методология) в городах и сельских районах различны из-за разли-

чий в характере экономической активности. Иногда страны проводят отдельные обследования в городских и сельских местностях на основе различных подходов к организации работы (периодов обследования и др.). Можно, однако, утверждать, что при организации обследований рабочей силы многие развивающиеся страны уделяли *недостаточно внимания различиям в требованиях к данным и условиям обследования в городских и сельских районах.*

Во-вторых, исключение некоторых районов и групп населения может диктоваться соображениями, связанными с затратами и другими практическими моментами. Следует периодически анализировать, сохраняют ли силу соображения, по которым были исключены некоторые районы и группы, и пытаться обосновать каждый случай такого исключения. Иногда представляется возможным предусмотреть для более сложных и дорогостоящих случаев пропорционально меньшие объемы выборки, а не исключать их полностью из обследования. Важно также документально оформить в каждом случае масштабы исключения из обследования, так как от них зависят общие выводы, которые научно оправданы на основе результатов выборочного обследования; насколько это возможно, следует оценить влияние исключений на степень представительности результатов для всей страны. Иногда сравнительно небольшие исключения могут и не повлиять существенно на отношения и коэффициенты, исчисленные на основе обследований, но могут более серьезно отразиться на оценках агрегированных показателей для всего населения.

Временной аспект

Единицы и их характеристики со временем меняются в любой совокупности. Помимо отбора единиц из совокупности обследование предполагает также выбор из общего интервала определенного отрезка времени, к которому относятся результаты выборочного обследования. Представительности в пространственном смысле часто справедливо отдают предпочтение перед представительностью во временном отношении, так как корреляция показателей для данной единицы во времени выше, чем для соседствующих в пространстве единиц (Kish, 1965, section 12.5). По этой же причине временные рамки допускают большую неопределенность по сравнению с географическими рамками, которые необходимо устанавливать более строго.

Тем не менее четкое определение временных рамок может иметь весьма важное значение, особенно когда преобладают явно выраженные сезонные, циклические и иные колебания показателей во времени. Статистические выводы на основе выборочных наблюде-

ний средних условий в период обследования можно сделать лишь в том случае, если имеется представительный выбор времени (отдельные месяцы, кварталы, сезоны и т.п.) из общего периода (например года). Указанное требование на практике часто необоснованно игнорируют. Соображения относительно планирования выборки с учетом этого требования более подробно рассмотрены в парагр. 5.

3. ОСНОВА ВЫБОРКИ И ИСПОЛЬЗОВАНИЕ ГЛАВНОЙ ВЫБОРКИ

Как только обследуемая совокупность определена, ее следует представить в физической форме, на базе которой можно было бы осуществить отбор требуемого типа. *Основы выборки* и служат этим целям. Она должна обеспечить отбор "вероятной" выборки элементарных единиц, образующих совокупность, т.е. отбор, при котором использование соответствующих рандомизированных (случайных) процедур гарантирует, что каждая единица имеет известную, не нулевую вероятность быть отобранной.

Проблема основы и планирования выборки для программ обследования домашних хозяйств подробно рассмотрена в работе ООН (United Nations, 1986). Настоящий параграф посвящен некоторым важным моментам, представляющим особый интерес с точки зрения исследований рабочей силы, и разъяснению ряда пунктов.

Отбор выборки для обследований домашних хозяйств обычно осуществляют в несколько ступеней, на первой отбирают первичные единицы выборки, которые, как правило, представляют собой территориальные единицы сравнительно крупных размеров. Указанные единицы должны иметь четкие непересекающиеся границы и совместно должны исчерпывающим образом охватывать изучаемую совокупность. Основы выборки будут состоять из полного перечня территориальных единиц с картами, описаниями и другой информацией, относящейся к их идентификации и отбору. Хорошие основы выборки включают также информацию о размерах и иных характеристиках единиц, что улучшает качество отбора и оценки. Основы из первичных единиц выборки, охватывающая всю обследуемую совокупность, именуется *первичной основой выборки*.

При многоступенчатой процедуре отбор на каждой ступени ограничен единицами, отобранными на предшествующих ступенях. Сказанное означает, что коль скоро на базе первичной основы выборки осуществлен отбор первичных единиц, основы для последующей ступени отбора необходимы лишь применительно к отобранным первичным единицам. Это же относится к отбору на второй ступени и т.д. В иерархии основ все основы, следующие после первичной, именуется *вторичными основами выборки*.

Первичные основы выборки

Между первичной и вторичными основами выборки существует принципиальное различие. Так как вторичные основы ограничены элементами, отобранными на предшествующих ступенях, часто представляется возможным составлять и сохранять их специально для одного обследования или ограниченного отрезка времени. Первичная же основа выборки необходима для охвата всей совокупности. Создание и поддержание полной и надежной первичной основы выборки, включающей элементы нужного вида и размера, наряду с необходимой вспомогательной информацией является решающим условием проведения обследований рабочей силы и иных широко-масштабных работ. Первичная основа выборки — это крупная инвестиция, которую обычно осуществляют в расчете на неоднократное использование для разных обследований и их раундов в течение некоторого периода времени. Указанные условия обычно выполняются применительно к национальным обследованиям рабочей силы в связи с их, как правило, крупными масштабами, широким охватом и жесткими требованиями, а также непрерывным характером работ. Кроме того, как отмечалось выше, обследования рабочей силы хорошо подходят для планирования и организационной увязки с другими обследованиями рабочей силы, так что затраты по разработке и поддержанию хороших первичных основ выборки часто поддаются распределению между обследованиями рабочей силы и иными обследованиями.

Требуемые свойства хорошей основы выборки можно объединить в три группы (United Nations, 1986, p. 63—79): 1) свойства, связанные с качеством и позволяющие минимизировать ошибки регистрации, включая наличие четко очерченных, не устаревших и стабильных элементов с соответствующими определителями и полный охват совокупности; 2) свойства, относящиеся к эффективности и обеспечивающие эффективное проектирование и оценки, включая наличие элементов с хорошим описанием и последней статистической информацией об их размерах и других характеристиках для стратификации (разбиения) и отбора; 3) свойства, относящиеся к затратам и обеспечивающие экономическое сопоставление, использование и ведение основы выборки.

Основные этапы разработки и ведения первичных основ выборки рассмотрены ниже. Они включают: определение целей и стратегию максимизации полезности основы выборки для различных задач; выбор подходящего элемента для основы; физическое представление этих элементов для облегчения доступа к ним и их использова-

ния; введение и обновление основы с целью ее повторного использования на протяжении некоторого времени.

Определение целей и стратегии

Оправданная величина инвестиции в разработку и ведение первичной основы выборки зависит от того, насколько интенсивно и долго ее будут использовать. Основа выборки может служить рамками для отбора связанных и не связанных (независимых) выборок для разнообразных обследований домашних хозяйств. Ее можно разработать с учетом административного деления страны и увязать с банками статистических данных различной степени агрегирования из разных источников. В частности, удачная стратегия заключается в тесной увязке разработки первичной основы выборки с работами в области переписи населения, а при планировании переписи — в учете требований выборки, включая отбор счетных районов и участков, систему нумерации, планы табулирования данных по счетным участкам, картографическое отображение результатов и т.п.

Разумеется, трудно рассчитывать на то, что одна первичная основа выборки будет удовлетворять потребности всех возможных обследований. Могут оказаться необходимыми отдельные основы выборки для обследований, требующих иных элементов, например, обследований предприятий, районных сельскохозяйственных обследований или обследований, предполагающих доступ к некоторым специфическим группам населения. Но и в этих разнообразных ситуациях иногда может пригодиться, и для разных целей, удачная основа выборки, относящаяся ко всему населению. Во-первых, путем реорганизации и перегруппировки различных показателей размера и других характеристик элементов основы выборки можно одну и ту же основу приспособить для эффективного отбора элементов, требуемых для различных обследований. К примеру, в рамках Национального выборочного обследования Индии прибегают к известной перегруппировке первичной основы выборки, базирующейся на счетных участках переписи, с учетом конкретных целей каждого раунда обследования (Murthy and Roy, с. 15—16). Во-вторых, основу общего назначения можно иногда сделать пригодной для более специальных целей путем включения в нее некоторой дополнительной информации. Например, Мерти и Рой (1970) описывают обследование небольших промышленных предприятий, проведенное с помощью базирующейся на переписи основы выборки с включением в нее дополнительной информации о концентрации

предприятий, требовавшейся для обследования. (Индонезия прибегла к аналогичному подходу в переписи населения 1990 г.). Приведенный конкретный случай служит также хорошим примером из практики развивающейся страны, который показывает, как качество основы выборки можно улучшить путем привлечения информации из нескольких источников. В-третьих, общую основу можно использовать для пополнения перечней или других видов основ специального назначения. Основы специального назначения или передачи часто содержат информацию об элементах, представляющих особый интерес (в частности, единиц крупного размера), что обеспечивает эффективный отбор, но не полный охват элементов, представляющих особый интерес, в рамках всей совокупности. Хотя основа общего назначения не является лучшим источником для отбора элементов в этих случаях, она может служить необходимым дополнением, обеспечивающим полный охват.

Выбор типа первичных элементов выборки

Весьма важное решение при разработке и составлении основы выборки заключается в определении первичных элементов выборки. Этот выбор диктуется различными соображениями: а) возможным использованием первичной основы выборки; б) имеющимися источниками и материалами для создания основы, включая связанные с этим затраты и усилия; в) проблемами, связанными с получением данных, проектированием и применением выборки, включая издержки и требования контроля. Цель должна заключаться в том, чтобы обеспечить полноту охвата, гибкость при непредвиденных вариантах использования основы выборки, стабильность элементов при повторном длительном использовании, возможность увязки смежных обследований и их раундов, базирующихся на одной основе, простоту ведения и обновления и, разумеется, статистическую эффективность проектирования и оценки.

Выбор первичных элементов выборки может существенно повлиять на качество и стоимость обследования. Элементы должны быть хорошо определены, четко очерчены, иметь хорошее описание для выявления и демаркации, содержать самую последнюю информацию о размере и других характеристиках. Элементы должны быть стабильными во времени, особенно когда обследование предполагает неоднократный счет или распространяется на продолжительный период либо существует значительный разрыв во времени между составлением основы и отбором, а также проведением обследования. Желательными свойствами основы служат также низкие затраты и

приемлемые усилия при получении, ведении, использовании и обновлении элементов выборки. Первичные элементы должны быть “соответствующего” размера с учетом организации и структуры издержек, связанных с операцией по сбору данных обследования; они должны также быть достаточно единообразными по размеру в рамках каждой выборочной области. Если элементы слишком велики, то отбор и обработка достаточного их количество, для получения надежной выборки могут оказаться невозможными; кроме того, издержки составления и ведения вторичных основ выборки, равно как издержки по сбору данных при крупных первичных элементах в ряде случаев могут оказаться чрезмерно высокими. С другой стороны, если элементы слишком малы, их границы могут оказаться недостаточно четкими, а сами они недостаточно стабильными во времени. Может представиться затруднительным получить достаточно представительную выборку *внутри* первичных элементов или удовлетворить нужды различных обследований и их раундов, требующих отдельных выборок из числа тех же элементов.

“Соответствующий” размер элементов — это слишком сложный вопрос, чтобы на него можно было дать общий ответ, ибо он зависит от конкретной ситуации. Однако с учетом практики проведения обследований рабочей силы и других обследований домашних хозяйств можно достаточно отчетливо выделить четыре варианта.

Первый вариант — один из крайних полюсов — используются весьма крупные первичные элементы выборки. Общее число элементов совокупности обычно невелико, и для выборки из них отбирают небольшое количество. Эти элементы должны быть относительно разнородны и тщательно оформлены. Каждый из них требует значительных затрат на подготовку карт, списков и вторичных основ выборки, а также, что, пожалуй, еще важнее, на подбор, обучение и размещение счетчиков в выборочных районах. Если для перемещения счетчиков между вторичными элементами поездки могут быть минимальными или вообще не требоваться, то поездки внутри этих крупных единиц могут быть довольно значительными и требовать наличия дешевого и удобного местного транспорта. Другой аспект такого подхода состоит в том, что для сохранения существенных инвестиций, связанных с определением элементов выборки, целесообразно использовать одно и то же множество элементов для различных обследований и их раундов в течение определенного времени. “Ротацию”, обновление выборки придется осуществлять преимущественно в рамках фиксированного множества первичных элементов; даже на этапе полной реорганизации выборки может представиться желательным сохранение в новой выборке как можно

больше элементов, насколько это статистически допустимо. При правильной постановке дела основные преимущества этой системы могут включать: стабильность границ и других характеристик крупных элементов; экономия и удобства, связанные с продолжительным использованием одного и того же множества выборочных районов; уменьшение потребности в поездках счётчиков. Несмотря на небольшое число отобранных элементов, выборка может оказаться достаточно эффективной благодаря крупному размеру и разновидности элементов.

Наиболее известным и документально оформленным примером такого использования первичных элементов выборки, пожалуй, служит текущее обследование населения в США, представляющее собой крупномасштабное ежемесячное обследование, цель которого “обеспечить оценки занятости, безработицы и других характеристик общей рабочей силы, населения в целом и его различных подгрупп” (United States, 1978). В этом обследовании крупнейшие элементы, включенные в процессе отбора (так называемые “несамостоятельные первичные элементы отбора”), могут достигать 4000—5000 кв. км и иметь до четверти миллиона жителей. Система позволяет набирать счётчиков на неполное рабочее время из числа местных жителей. Часто это домохозяйки, которые при соответствующей подготовке и опыте, как правило, обеспечивают высокое качество работы, несмотря на небольшую оплату. Счётчики обычно работают, проживая у себя дома, и покрывают значительные расстояния в пределах своего участка на собственном автомобиле.

Второй вариант — второй полюс — выборки состоят из многочисленных очень небольших первичных элементов. Такая система может подходить для густонаселенных городских районов, где имеются списки жилых единиц и/или весьма подробные карты небольших участков района, а благодаря хорошо развитой транспортной сети и коротким расстояниям поездки между элементами не представляют серьезной проблемы. Иллюстрацией этой схемы опять-таки может служить текущее обследование населения в США. В этом обследовании в типических группах городов (так называемых “самостоятельных” первичных элементах отбора) непосредственно отбирают небольшие участки, именуемые сегментами (фактически они образуют элементы первичного отбора). Количество таких отобранных элементов достаточно велико, так как в любой данный момент каждый такой отобранный элемент пополняет общую выборку лишь на несколько (четыре или примерно столько) жилых единиц.

Ни один из указанных двух крайних вариантов не получил распространения в развивающихся странах, так как в общем главные условия для их эффективного применения здесь не удовлетворяются. Однако, как будет показано ниже, в некоторых развивающихся странах весьма крупные первичные элементы используются, но обычно по совершенно иным соображениям. Что же касается использования в крупных элементах местных интервьюеров, которые были бы весьма мобильны в пределах выборочного района, то следует заметить, что во многих развивающихся странах подобрать на месте квалифицированных и опытных людей трудно. Наблюдение и контроль при наличии постоянного персонала, занятого полное рабочее время, как правило, обеспечить легче, чем при использовании случайных работников, занятых неполное время. Кроме того, условия передвижения в развивающихся странах могут быть совершенно иными: часто здесь может существовать достаточно развитая сеть магистральных дорог, облегчающая поездки между районами, тогда как местные поездки могут оказаться более затруднительными и утомительными даже на небольшие расстояния (Verma, 1977).

Редко прибегают в развивающихся странах и к небольшим элементам первичного отбора, так как отсутствуют карты и другие материалы для определения небольших территориальных элементов. Беспорядочная застройка и отсутствие должной адресной системы также препятствуют этому. Тем не менее в развивающихся странах, особенно в городах, проводили обследования на базе многочисленных и небольших элементов первичного отбора. Примерами могут служить обследования рождаемости в Венесуэле, Колумбии, Панаме, Перу, Филиппинах и Шри-Ланке (Verma, 1980).

Третий вариант (преобладающая система в большинстве развивающихся стран) состоит в использовании основы первичного отбора для обследований рабочей силы и других обследований домашних хозяйств на базе счетных участков последней переписи населения. Как правило, такие участки включают в среднем от нескольких десятков до нескольких сотен домашних хозяйств. Во многих случаях счетные участки переписей достаточно единообразны по размеру и располагают вполне удовлетворительными картами, описаниями и информацией о численности населения и других необходимых характеристиках. Обычно счетные участки определяют как подразделения административных и других крупных районов, их можно расположить в иерархическом или географическом порядке, чтобы найти типические группы и отчетные единицы, либо объединить их для образования более крупных элементов для основы выборки.

Иногда основу первичного отбора можно также улучшить путем раздробления некоторых крупных счетных участков и объединения более мелких в элементы соответствующего размера.

Четвертый вариант, заслуживающий упоминания, заключается в практике использования в ряде развивающихся стран немногочисленных и весьма крупных элементов первичного отбора. На это может существовать несколько причин, но все они, как правило, отличны от соображений, рассмотренных выше применительно к первому варианту. В некоторых странах пригодные первичные основы выборки с небольшими (и, как правило, более подходящими) элементами попросту отсутствуют, по меньшей мере для части населения. Обычно это отражает неудовлетворительную картографическую работу во время переписи. Например, во многих переписях в Камеруне в качестве первичных элементов использовали *arrondissements* (районы), которых в стране только 144. При обследованиях во многих арабских странах, например, Иордании, Сирии и Египте, первичными элементами считали целые города. Отчасти это может быть связано с отсутствием подходящей основы для меньших районов, а отчасти — с затратами и практическими трудностями распространения выборок на много разных городов. Вторая главная причина использования весьма крупных первичных элементов в ряде развивающихся стран заключается в действительных или мнимых административных удобствах. Основы для отбора намного меньших районов часто уже существуют или их можно разработать, то эту информацию используют лишь для создания основ вторичного отбора в рамках элементов первичного отбора. В качестве последних, например, при обследованиях рабочей силы и других обследованиях в Таиланде до последнего времени прибегали к “чангват” (провинциям), которых в стране только 71. И в Китае при последних переписях использовали многоступенчатый отбор, причем первичными элементами выступали округа (средняя численность населения — полмиллиона). В обоих этих случаях причины носили преимущественно административный характер, а не были связаны с наличием основ для отбора более мелких единиц или тщательным анализом издержек по сбору данных.

Физическое представление основы выборки

Физическое представление элементов должно отвечать общей цели обеспечения удобного доступа, использования и работы с элементами. Еще одно соображение — это получение сводных данных и контрольных итогов, которые помогают при планировании выборки

и оценке, а также обеспечивают средство для проверки репрезентативности и полноты основы выборки. Такая первичная основа должна представлять особый упорядоченный список элементов в соответствии с административным делением, основными отчетными объектами, типическими и другими содержательными группировками, которые сами должны быть упорядочены с учетом расположения, размера и других полезных критериев. Важно, чтобы каждый элемент располагал *единственным определителем* и информацией, используемой для классификации и упорядочения, которые могут быть различными в зависимости от характера использования основы выборки. В этом отношении может оказаться особенно целесообразной компьютеризация основы выборки. В нее можно также включить сведения об использовании каждого элемента в прежних выборках, обобщенные характеристики элементов, карты и описания для определения и выявления элементов.

Процедуры ведения и обновления основы

Прежде всего необходимо обеспечить разработку планов и процедур ведения и периодического обновления основы для отбора элементов, а также чтобы эти процедуры были должным образом документально оформлены, в работе ООН (1986, с. 102) содержатся отрицательные примеры стран, где этого, похоже, не было сделано.

Целесообразно как можно последовательнее централизовать работу и стандартизировать процедуры, связанные с обновлением. Это особенно важно для крупных статистических органов, когда к той же основе отбора прибегают многие пользователи и в различных целях. В то же время статистические органы должны сотрудничать с другими национальными организациями, специализирующимися на выпуске карт различного назначения, и опираться на их продукцию.

Как правило, удобнее всего вносить крупные изменения в основы отбора в связи с переписями населения. Более ограниченные коррективы можно вносить периодически между переписями. Такое обновление предполагает выявление соответствующих источников информации, определение правильной частоты обновления и критериев внесения в основу дополнительных элементов и исключения несуществующих единиц, а также установление процедур регистрации изменений границ и характеристик элементов. Важно рассмотреть, следует ли соотнести старые элементы и информацию с новыми. Обновление основы часто может означать ни включение,

ни исключение элементов, а лишь изменение определений, границ и характеристик существующих списков (элементов).

Наконец, можно отметить, что не все элементы должны обновляться одинаковым образом и в одно время. Обновление нередко приносит пользу, когда можно охватить им лишь часть элементов.

Основы вторичного отбора

При многоступенчатом плане выборки основа для дальнейшего отбора на каждой ступени необходима лишь применительно к элементам, выбранным на предшествующей ступени. По указанной причине иногда целесообразно разработать и вести основы вторичного отбора для определенного обследования и использования в течение более ограниченного периода времени по сравнению с основами первичного отбора. В то же время основы вторичного отбора, как правило, менее долговечны, так как в общем элементы и их характеристики менее стабильны по мере перехода от крупных элементов более высокого порядка к элементам меньших размеров и более низкого порядка.

Несмотря на эти оговорки, разработка и ведение основ вторичного отбора могут быть сопряжены с весьма значительными затратами, что предполагает их неоднократное и продолжительное использование, насколько это позволяет стабильность элементов и их характеристик. При планировании обследования, где элементы первичного отбора довольно велики и пространны, выборка часто охватывает также значительную часть совокупного населения, поэтому требуются разработка и ведение основы отбора элементов на второй ступени. Основа вторичного отбора внутри элементов первичной выборки может быть связана с такими затратами, что становится крайне желательным сохранить постоянную выборку элементов для повторного использования в течение ряда лет и максимальное число старых элементов даже при основательной ревизии выборки, после новой переписи населения. В указанном контексте уже упоминался пример текущего обследования населения в США. Это справедливо и в отношении ряда других обследований, например, обследований рабочей силы в Канаде.

При планировании обследований на базе более мелких и компактных элементов первичного отбора затраты по разработке и ведению основ вторичного отбора в расчете на элемент могут быть ниже. Но в этом случае число элементов, как правило, намного больше, так что суммарная стоимость основ вторичного отбора может также представлять собой крупную инвестицию.

Элементы вторичного отбора на двух последних ступенях представляют особый интерес, а потому рассмотрены ниже более подробно. Кратко охарактеризованы также правила объединения различных видов элементов при обследованиях рабочей силы.

Конечные территориальные единицы

Конечные территориальные единицы — это *территориальные* единицы на последней ступени отбора. Часто это относительно небольшие и компактные элементы, после которых процесс отбора переходит от географического выбора к отбору из списков лиц, домашних хозяйств, жилых единиц или их небольших групп (гнезд). Связано это с тем, что по мере уменьшения территориальных единиц границы между ними становятся все менее четкими и более трудноопределимыми. После определенного момента дальнейшее деление на более мелкие территориальные единицы становится более дорогостоящим и менее эффективным по сравнению с непосредственным перечислением и отбором конечных единиц. Еще одна причина, по которой после определенного момента переходят к этому методу, состоит в том, чтобы обеспечить распространение выборки конечных элементов внутри гнезд. Во многих развивающихся странах, особенно в сельской местности, отсутствие хороших карт и удовлетворительной адресной системы, а также, как правило, беспорядочная застройка ведут к тому, что конечные территориальные единицы должны быть довольно большими по размеру, часто это целые деревни. В этих случаях конечные территориальные единицы могут служить базовыми элементами для организации составления списков и сбора данных, а также ведения отчетности о выборке и ее использовании. На практике во многих обследованиях конечные территориальные единицы настолько крупные, что весь план может быть связан с двумя ступенями: конечные территориальные единицы сами являются элементами первичного отбора, а внутри отобранных конечных территориальных единиц составляют списки и осуществляют выбор отдельных домашних хозяйств или аналогичных элементов выборки (см. например, Verma, 1977). Индивиды, проживающие в той же местности или поблизости, могут обладать схожими характеристиками и одинаковой экономической активностью. Поэтому при обследованиях рабочей силы, как правило, необходимо охватить выборкой много конечных территориальных единиц, избежать использования очень маленьких и компактных единиц, а в каждой конечной территориальной единице отбирать лишь небольшое количество индивидов.

Конечные единицы отбора

Конечные единицы отбора — это окончательные единицы, полученные в конце процесса отбора; они либо идентичны элементам населения, которые служат единицами наблюдения и анализа, либо связаны с этими элементами. При обследованиях рабочей силы, как и в других обследованиях домашних хозяйств, конечными единицами отбора могут быть: а) отдельные лица; б) домашние хозяйства; в) адреса или жилые единицы; г) небольшие группы соседних жилых единиц, определенные на основе списков и эскизных карт; д) сами конечные территориальные единицы, внутри которых пронумерован каждый элемент. В приведенном перечне единицы расположены в порядке возрастания степени постоянства или стабильности соответствующих элементов.

Вариант “д” именуют отбором “бери все” или “компактным гнездом”. При этом отсутствуют какие-либо подвыборки, а возможно, и отдельные списки домашних хозяйств или домов внутри конечных территориальных единиц. Преимущества такого подхода заключаются в большей стабильности единиц и меньшем числе ошибок, связанных с охватом. Однако он не очень подходит для обследований рабочей силы, особенно в развивающихся странах. Общая проблема в работах по обследованию, связанная с отсутствием карт для небольших единиц и других картографических материалов, заключается в том, как избежать слишком больших территориальных единиц, требующих значительных затрат при составлении списков и сборе данных. Кроме того, территориальные единицы, которые существуют или могут быть образованы, часто недостаточно однородны по размеру, так что эффективный отбор предполагает использование схемы типа “вероятность, пропорциональная размеру”, требующей последующей подвыборки. Как уже отмечалось, лица, проживающие на одной территории, могут обладать схожими характеристиками в сфере экономической активности, так что предпочтительнее выбрать небольшое число индивидов из многих территорий, нежели многих индивидов из небольшого числа территорий. Однако при обследованиях, направленных на выявление редких элементов в совокупности, подход “бери все” или “компактных гнезд” может быть удачной схемой отбора (Kish, 1965, section 11.4).

Отбор небольших групп из соседних единиц по спискам (вариант “г”) позволяет избежать чрезмерного гнездования конечных единиц, присущего варианту “д”, и в то же время сохранить некоторые преимущества — большую стабильность единиц и меньшее число

ошибок, связанных с охватом. (Обсуждение возможных достоинств этого варианта в отличие от индивидуального отбора жилых единиц или домашних хозяйств, с точки зрения уменьшения числа таких ошибок см. в работе Киша и Хесса, 1958 г.) Тем не менее частых упоминаний о применении этого варианта в обследованиях в развивающихся странах не было; видимо, было бы целесообразно использовать его чаще, когда это возможно.

Вариант "а", т.е. составление списков и непосредственный отбор индивидов, может оказаться необходимым в определенных особых условиях. Так, в Сенегале и некоторых других странах многие люди живут большими общинами, где трудно или неудобно выделять в качестве единиц для отбора домашние хозяйства или жилые единицы, так что единственный практический выход — это составить список и отобрать индивидов непосредственно из числа жителей общины. Указанная схема подходит также, когда желательно минимизировать случаи отбора более одного индивида из одного домашнего хозяйства. Но вообще говоря, составление списков и отбор отдельных лиц в качестве конечных единиц — это излишняя и дорогостоящая процедура и к ней следует прибегать при минимальном разрыве во времени в связи с быстрыми изменениями обстановки.

Поэтому во многих обследованиях выбирают между вариантами "б" и "в", т.е. между домашними хозяйствами и жилыми единицами. Последние — это структурные единицы, определяемые по их адресу и местоположению, тогда как домашние хозяйства — это социальные образования, определяемые по входящим в них лицам. Составление списков жилых единиц обычно не требует контактов с индивидами или домашними хозяйствами и поэтому это можно сделать быстрее и с меньшими затратами, нежели составление списков домашних хозяйств. Те домашние хозяйства, которые нельзя сразу выявить на основе внешних признаков, как правило, требуют дополнительного уточнения. Но еще важнее тот факт, что, когда списки предполагается использовать на протяжении продолжительного периода или неоднократно, как это имеет место при непрерывном обследовании рабочей силы, жилые единицы обладают преимуществом большего постоянства по сравнению с домашними хозяйствами (United Nations, 1986, p. 113—114).

Существуют, однако, ситуации, когда использование в качестве конечных элементов отбора домашних хозяйств, а не жилых единиц, становится более предпочтительным или неизбежным. Информация о характеристиках домашних хозяйств может, к примеру, обеспечивать более последовательное расслоение элементов, что

важно с точки зрения эффективности выборки (хотя простое, но полезное расслоение элементов возможно и на базе списков жилых единиц; пример см. у Киша и Хесса, 1958). Во всяком случае преобладающие условия жизни и отсутствие нормальной адресной системы в сельской местности многих развивающихся стран приводят к тому, что содержательному определению и идентификации в большей мере поддаются домашние хозяйства, чем жилые единицы.

В качестве конечных элементов отбора жилые единицы чаще использовались в обследованиях Латинской Америки, а домашние хозяйства — в Азии и Африке. Представляется вероятным, что в целом при обследованиях рабочей силы более стабильные жилые единицы (или даже небольшие гнезда последних) обладают некоторыми преимуществами, и к ним следует прибегать, когда обстоятельства позволяют делать это.

Правила объединения

Обследование может потребовать учета единиц разного типа. Во-первых, имеются элементарные единицы (например индивиды), образующие и определяющие изучаемую совокупность. Единицы, по которым *собирают* информацию, и единицы, применительно к которым ее *анализируют*, могут отличаться от элементарных единиц, а иногда друг от друга. Обследование домашних хозяйств может быть рассчитано на получение и анализ информации, относящейся к нескольким из этих типов. Например, если при обследовании рабочей силы собирают информацию преимущественно по отдельным лицам, одновременно можно собирать данные и о доходе, жилищных условиях и других социально-экономических характеристиках, относящихся к семьям, домашним хозяйствам, общинам и другим социальным образованиям. Кроме того, данные, собранные на более низких уровнях, можно агрегировать и анализировать применительно к единицам более высокого уровня или, наоборот, данные для единиц более высокого уровня можно отнести к каждому более низкому уровню и соответственно анализировать их. (См. работу Киша, 1965а, где содержится интересное обсуждение некоторых вопросов множественности уровней сбора и анализа информации; более обстоятельное обсуждение в контексте данных на уровне общины имеется в различных статьях сборника под редакцией Кэстерлина, 1985).

Во-вторых, в многоступенчатой схеме отбора присутствуют элементы различного типа. Как уже отмечалось, обычно речь идет об иерархии территориальных единиц с жилыми единицами и домаш-

ними хозяйствами на последней ступени отбора. Элементы отбора могут совпадать с элементарными единицами сбора и анализа информации.

В-третьих, часто существует различие между названными выше типами единиц и единицами, которые фактически предоставляют информацию. К примеру, информацию о домашних хозяйствах могут предоставить любой взрослый житель или специально назначенное лицо; точно так же информацию об индивиде он может предоставить сам или же это могут сделать за него другие лица. Единицы, предоставляющие информацию, могут быть определены в соответствии с правилами получения ответов, установленными для обследования.

В связи с наличием при обследовании многообразия типов единиц необходимо установить соответствующие *правила объединения* единиц. Цель названных правил — обеспечить *вероятностную* выборку, внутри которой каждая элементарная единица, а следовательно, каждая аналитическая единица совокупности имела бы известную ненулевую вероятность быть учтенной при обследовании. Правила объединения требуются для: а) различных уровней единиц отбора; б) конечных единиц отбора и элементарных единиц; в) элементарных единиц и единиц сбора и анализа информации; г) единиц сбора информации (к которым информация относится) и респондентов обследования (которые предоставляют информацию).

Использование главной выборки

При любом обследовании с многоступенчатым отбором каждая ступень процесса отбора предполагает подготовку основы для отбора и выборку вплоть до получения выборки из конечных единиц отбора. В целях экономии и удобства одну или более ступеней этой работы можно объединить с другими обследованиями. Выборка, полученная в результате объединения ступеней ряда обследований, именуется *главной выборкой*. Ее цель — обеспечить общую выборку элементов до определенной ступени, после которой дальнейший отбор можно производить с учетом нужд конкретного обследования.

Например, после подготовки основы для первичного отбора можно отобрать выборку первичных элементов для использования в нескольких обследованиях или их раундах. Все или некоторые элементы первичного отбора можно затем взять из главной выборки для продолжения процесса отбора с учетом требований определенного обследования. В этом случае задача построения и разбиения первичной основы и отбора первоначальной выборки одинакова для всех

обследований, опирающихся на главную выборку. На основе первоначально отобранных первичных элементов *можно* осуществить отбор подвыборки для отдельных обследований. Хотя подвыборки из первичных элементов для различных обследований базируются на главной выборке, они могут быть независимыми или связанными, идентичными, частично пересекающимися или совершенно разными.

Первоначальный отбор первичных элементов, пригодных для многократного использования, осуществляется на основе *минимального* требования, предъявляемого к концепции главной выборки. Разумеется, нередко целесообразно довести процесс совместного отбора до более низких ступеней, с тем чтобы получить главную выборку элементов второй ступени или конечных территориальных единиц либо даже конечных единиц отбора. В этом случае отбор для конкретных обследований будет осуществляться на самой последней ступени, применительно к единицам этой главной выборки.

Цели использования концепции главной выборки заключаются в следующем:

а) экономия затрат на разработку основ и других материалов для отбора, а также на сам отбор путем совмещения выборок для нескольких обследований;

б) облегчение содержательной и организационной увязки различных обследований и их раундов;

в) ускорение (равно как ограничение и контроль при необходимости) извлечения многих выборок для различных обследований на базе одной основы отбора элементов.

Последний пункт может представлять особую ценность, когда в проведении обследований одной совокупности заняты несколько разных организаций.

Все это нередко веские причины для использования концепции главной выборки при проведении обследований рабочей силы. Во-первых, как будет показано подробнее в парагр. 5, обследование рабочей силы часто проводят раундами или более мелкими этапами, каждый из которых предполагает учет пространственно репрезентативной выборки в течение определенного времени. Эти выборки нельзя отбирать независимо, они должны соотноситься определенным образом в зависимости от содержательных и организационных требований обследования. Их отбор на базе общей главной выборки облегчает такую увязку.

Другая причина в том, что обследования рабочей силы, представляющие собой крупномасштабные непрерывные работы, часто увязывают тем или иным образом с обследованиями по другим проблемам (см. гл. 10, парагр. 3). В частности, как только налажено

крупное обследование рабочей силы, его можно удачно использовать для поддержки обследований домашних хозяйств по другим направлениям. Такие возможности следует расширять, но одновременно и контролировать, с тем чтобы вспомогательные операции отрицательно не отразились на основном обследовании рабочей силы, например, в результате совпадения времени проведения или чрезмерной нагрузки на респондентов из-за многократного использования одной и той же выборки. Главная выборка — это инструмент налаживания и в то же время контроля, согласования и увязки работ.

4. НЕКОТОРЫЕ ПРАКТИЧЕСКИЕ АСПЕКТЫ ПЛАНИРОВАНИЯ ВЫБОРКИ

Планирование выборок для комплексных социально-экономических обследований — это весьма специальная задача, по которой имеется обширная литература теоретического и практического характера. Охватить этот вопрос полностью в настоящем руководстве не представляется ни необходимым, ни возможным. Цель последующих разделов — лишь дать несколько советов, которые могут оказаться полезными в практической работе по планированию выборки для обследований рабочей силы.

Необходимость практической ориентации

Обследования рабочей силы, как правило, представляют собой крупномасштабное мероприятие, преследуют несколько целей и опираются на сложные методы отбора. При этом желательно опираться на надежные и практические методы и процедуры, с тем чтобы фактически проведенные операции по сбору и обработке данных не слишком отклонились от плана. Эти операции многочисленны, они включают: разработку плана, выбор конторы, наблюдения, кодирование, вычислительные работы и статистический анализ. Чтобы завершить цепь этих операций с известной надежностью, процедуры отбора должны быть достаточно “ошибкоустойчивы” (Kish, 1977). Предпочтительнее планы и процедуры, которые, будучи не самыми точными и тонкими в оптимальных условиях, могут тем не менее устоять в неожиданных и незнакомых ситуациях, неизбежных в практике обследований, и особенно — обусловленных ошибками человека. Практический отбор должен целиком определяться характером и взаимосвязями операций по сбору и обработке данных, а также обследования в целом. Составитель плана обязан постоянно помнить, чего можно и чего нельзя ждать от счетчиков и конторских работников, занятых реализацией выбор-

ки. Хотя теория выборки обеспечивает основу для практического отбора, а нередко служит средством оценки произведенной выборки, она в большинстве случаев не способна дать единственное в своем роде или явно оптимальное решение. Стало почти трюизмом утверждение, что отбор элементов — это отчасти наука, отчасти искусство. Хотя это иногда повергает в ужас теоретиков, искусство находит себе место в науке, когда нужно “почувствовать” ситуацию, чтобы принять решение в условиях наличия ограничений на выбор вариантов. В конечном счете решающим фактором опять-таки является характер и взаимосвязь проводимых операций (Verma, 1977).

Применение вероятностной выборки

В парагр. 3 гл. 10 охарактеризованы некоторые общие черты обследований рабочей силы с учетом сложившейся практики в развивающихся и развитых странах. Было отмечено, что из-за нередко официального статуса и общенациональных масштабов обследований рабочей силы существует вероятность того, что к ним будут предъявлены довольно жесткие требования в отношении сроков, точности, внутренней непротиворечивости и сопоставимости во времени. Результаты обследования должны быть сопоставимы с данными из других источников. Кроме того, обследования рабочей силы сами могут быть использованы для получения различных данных по другим проблемам, равно как служить базой (в смысле предоставления выборки) для других обследований населения. Указанные требования выполняются только при наличии *вероятностных выборок* достаточно большого объема, полученных с помощью хорошей основы отбора, которая охватывает все население. Вероятностный отбор предполагает, что каждый элемент совокупности имеет известный ненулевой шанс быть отобранным для выборки, ибо лишь в этом случае результаты выборочного обследования можно объективно и научно обоснованно распространить на всю совокупность. При невероятностном отборе оценки для всей совокупности на основе выборки можно сделать, лишь прибегнув к некоторым дополнительным и, по сути, произвольным допущениям. Применительно к некоторым относительно небольшим обследованиям, призванным глубоко изучить некоторые взаимосвязи, невероятностный отбор можно признать приемлемым. Однако применительно к крупномасштабным обследованиям, особенно тем, которые должны дать оценки для всего населения, оправданий для отбора на основании квот, оценочных суждений и других невероятностных методов практически не существует.

Имеются, однако, ситуации, когда небольшие отступления от указанного общего принципа для обследований рабочей силы оправдать можно по соображениям экономии и практичности. Иногда из-за отсутствия соответствующей основы отбора и по другим причинам получение строго вероятностной выборки для небольшой части обследуемого населения невозможно или сопряжено с недопустимо большими затратами, поэтому приходится прибегать к приблизительным процедурам. Но любые отступления следует по возможности сводить к минимуму.

Некоторые методы отбора (например, выбор заранее определенных номеров из списков, в которых общее число номеров неизвестно) дают выборки, где относительная вероятность выбора элементов известна, но абсолютные значения вероятности выбора неизвестны. Такие выборки обеспечивают приемлемые оценки средних и других отношений (так как они зависят лишь от относительной вероятности и числа номеров в числителе и знаменателе), но не общие оценки для совокупности в целом, которые можно получить лишь с помощью внешней информации, если она имеется. По этой причине использования подобных методов отбора следует по возможности избегать.

Замена элементов, которые были отобраны для выборки и оказались вне ее рамок или не представляют интереса (например, пустующие дома), не только искажает вероятностный подход, но может сделать его недействительным. Уже только по указанной причине замена при обследованиях, как правило, не допускается.

Умеренность при определении объема выборки

Объем выборки должен, разумеется, быть достаточно большим с учетом содержательных и конкретных требований, предъявляемых к обследованию. Однако во многих случаях качество обследования пострадало из-за неоправданно большого объема выборки. Неоправданный выбор объема может быть следствием: 1) чрезмерного упора на точность выборки и забвения необходимости контролировать ошибки регистрации и обеспечивать своевременное получение результатов; 2) принятия неразумных и ненужных, с точки зрения целей обследования, стандартов точности; 3) стремления выдать слишком много группировок данных с чрезмерной их детализацией; 4) неспособности взвесить различие между содержательной и статистической значимостью результатов; 5) неспособности понять, что результаты нередко можно накапливать, с тем чтобы со временем обеспечить любую необходимую степень детализации; 6) нежелания уделить достаточно времени и средств для получения и

анализа информации, полезной для совершенствования процедур планирования и оценки; и прежде всего 7) недооценки расходов и усилий, требующихся для сбора и обработки данных.

Несмотря на названные практические соображения, выборка должна, конечно, быть достаточного объема, чтобы информация обладала точностью, необходимой для аналитических целей. Очевидно, что минимальный объем выборки зависит от вида анализа и характера использования данных. Основными определяющими факторами при этом являются: 1) количество отчетных единиц (например, географических районов или других подгрупп совокупности), по которым необходимо выдать отдельные таблицы или оценки; 2) степень содержательной детализации (группировки по отраслям, занятиям и т.п.) занятого населения; 3) тип и частота требуемых оценок (например, уровни за год или изменения по месяцам и т.п.); 4) степень требуемой точности отбора.

Требуемый общий объем выборки возрастает по мере увеличения количества отчетных единиц. Однако по ряду причин увеличение обычно не носит пропорционального характера. Во-первых, районы или подгруппы часто более однородны, чем население страны в целом, вследствие чего при той же точности отбора объем выборки для них может быть меньшим, чем для всей совокупности. Во-вторых, когда представляющие интерес подгруппы населения хорошо распределены по выборочной территории, тот же план выборки может быть относительно более эффективным, с точки зрения получения оценок, для таких подгрупп, чем для населения в целом. (Так называемый "эффект плана" для разработанной подгруппы ниже, чем для всей выборки) (см. гл. 14). В-третьих, требования к точности отбора на практике нередко менее жесткие применительно к отдельным территориям и группам, чем по стране в целом, так как оценки общенационального уровня считаются более важными в содержательном отношении.

Говоря о виде и частоте оценок, можно отметить, что на практике от любого обследования требуют ряда оценок различного типа, так что не существует одной оценки, которая полностью определяла бы объем выборки. Тем не менее различные оценки неодинаковы по степени важности, а следовательно, по требованиям к точности, что зависит от целей обследования; поэтому часто можно выделить подмножество оценок, которые можно считать определяющими с точки зрения минимального объема выборки. К примеру, требуемый объем выборки может быть значительно меньшим, когда главная цель — выявить *основные* категории занятого населения (скажем, распределение по основным отраслям, занятиям и категориям статуса),

чем в том случае, когда ставится задача проследить изменения в небольшой группе населения, например, безработных. Измерение изменений (особенно незаметных тенденций) обычно требует более крупных выборок по сравнению с измерением текущих уровней с той же точностью. Аналогичным образом выборочные обследования способны более точно оценивать уровни и отношения, нежели итоговые величины и абсолютные цифры для всего населения.

Отсюда вытекает, что не имеет смысла определять требования к объему выборки для обследований рабочей силы вообще. Чтобы дать некоторое представление о порядке связанных с этим величин, достаточно привести такой упрощенный пример. Предположим, главная задача обследования — дать оценку доли (например, уровня безработицы), составляющей примерно 10% определенной представляющей интерес подгруппы населения, и что ошибка не должна превышать 1% при высоком уровне значимости. В зависимости от эффективности плана выборки ее требуемый объем может составить порядка 5000 индивидов; при наличии в домашнем хозяйстве в среднем двух взрослых лиц выборка составит 2000 — 3000 домашних хозяйств. Чтобы при этом уровне точности получить оценки изменений или разностей, необходимый объем выборки может быть вдвое больше.

На практике процесс определения объема выборки может протекать примерно следующим образом. Зная основные области и подгруппы, для которых необходимо показать результаты отдельно, а также характер требуемых оценок, можно определить минимальный объем выборки, удовлетворяющий самым строгим требованиям к точности. Важно, однако, чтобы полученный таким путем объем выборки не превышал некий максимум, который считается целесообразным и практичным с учетом имеющихся ограничений. Если первоначальный объем превышает такой максимум, лучше пересмотреть и скорректировать цели и количество отчетных данных, а не пытаться навязать нереалистично большой объем выборки. При обследовании продолжающегося типа первоначальный отбор можно улучшить путем последующей оценки ошибок выборки и регистрации, с учетом изменившихся требований к анализу, а возможно, и практических ограничений.

Стратификация

Стратификация означает расслоение подлежащих отбору элементов на относительно однородные группы и независимый отбор в каждой группе. Можно отметить некоторые общие соображения,

относящиеся к стратификации. 1) Она часто способна уменьшить рассеяние выборки при минимальных дополнительных затратах. 2) Преимущества стратификации при простом случайном отборе часто весьма незначительны, однако они могут оказаться намного существеннее при многоступенчатом гнездовом отборе, к которому почти неизменно прибегают при крупномасштабных обследованиях домашних хозяйств. Поэтому обычно целесообразно посвятить значительные усилия для обеспечения хорошей стратификации, особенно для отбора элементов на первых ступенях. 3) Более правильно использовать несколько признаков стратификации с небольшим числом вариантов в каждом из них, нежели проводить группировку по одному признаку с множеством вариантов. Дело в том, что по мере увеличения числа “слоев” на основе единого признака преимущества стратификации уменьшаются. 4) “Слои” должны охватывать географические районы, для которых требуются оценки, или находиться внутри них (т.е. не выходить за их границы). 5) Во многих случаях простая географическая стратификация совместно со стратификацией по признаку типа населенного пункта являются наиболее действенным способом повысить эффективность и улучшить распределение выборки. Такую стратификацию просто осуществить, и она требует мало дополнительной информации. Она же, как правило, имеет многоцелевой характер и относительно стабильна во времени. 6) Систематический отбор на базе упорядоченных списков представляет собой экономичный и эффективный способ добиться преимуществ стратификации. 7) Нередко, используя специальные процедуры “контролируемого отбора”, целесообразно довести стратификацию до максимально возможной степени детализации, вплоть до момента, когда из слоя предстоит выбрать только один элемент, хотя для оценки ошибок выборки обычно из каждого слоя отбираются по меньшей мере два элемента. Часто предпочтительнее иметь выборку, которая более точна сама по себе, нежели менее точную выборку, но точность которой можно оценить более достоверно. 8) Важное назначение стратификации — обеспечить гибкость в выборе плана и процедур отбора. План и процедуры не обязательно должны быть одинаковыми для всех слоев.

От обследований рабочей силы, как правило, требуется охват всего населения, а не отдельных территорий или групп индивидов с определенными характеристиками. Поэтому самыми подходящими (и простейшими) признаками стратификации для отбора выборочных районов часто служат: а) тип населенного пункта, например, городской или сельский, или, что еще лучше, более детальная классификация, при которой различают населенные пункты различных

размеров — крупные города, более мелкие города, поселки, а также крупные, средние и небольшие деревни; б) административные и/или статистические регионы страны, предпочтительно с существенной детализацией на разных уровнях иерархической классификации; в) внутри каждого слоя — географическое упорядочение районов для систематического отбора. Если имеется информация об экономической активности, ее можно использовать для повышения эффективности отбора. Так, на базе хозяйственной переписи (или информации об экономической активности семей, полученной в ходе последней переписи населения), возможно, удастся разбить районы в зависимости от преобладающей отрасли или основного вида деятельности, обеспечивающей большую часть занятости в районе.

Внутри отобранных районов домашние хозяйства можно сгруппировать для последующего отбора в зависимости от характера информации, полученной на стадии составления списков домашних хозяйств (или жилых единиц). Полезным и широко используемым показателем является размер домашнего хозяйства. В городах составителям списков можно предложить дать хотя бы примерную и субъективную оценку социально-экономического статуса домашнего хозяйства (например, распределить их на три группы зажиточности — высокую, среднюю и низкую — на основе жилищных условий), что позволит осуществить стратификацию списков. Когда это необходимо, информацию можно собрать при составлении списков домашних хозяйств с целью выявить или отобрать домашние хозяйства с особыми характеристиками, которые могут представлять специальный интерес при обследовании (например, женщины как главы домашних хозяйств).

Самовзвешенный отбор

При обследованиях рабочей силы, равно как и при других обследованиях населения, существуют веские основания предпочесть “самовзвешенный” отбор. Такой отбор означает, что каждая элементарная единица совокупности имеет одинаковый ненулевой шанс попасть в выборку. Сказанное предполагает, что при единых правилах объединения элементарных единиц (например, индивидов) и конечных элементов отбора (например, домашних хозяйств, жилых единиц или их групп) конечные элементы отбирают с одинаковой постоянной вероятностью. На более высоких ступенях элементов можно, разумеется, отбирать с разной вероятностью, но так, чтобы эти различия на разных ступенях нейтрализовались. При

самовзвешенном отборе выборочные оценки можно получить без взвешивания данных, а их результаты затем при необходимости распространить на всю совокупность с помощью постоянного множителя.

Иногда в различных выборочных и/или отчетных областях необходимо прибегнуть к разным долям отбора, но даже в этом случае отбор в каждой области можно сделать самовзвешенным. Под *выборочной областью* понимают части совокупности, которые требуют различных планов и процедур отбора. Когда эти области существенно различаются по стоимости обследования их элемента или рассеянию элементов, задача “оптимального” распределения ресурсов может потребовать разных долей отбора в них. К примеру, стоимость обследования единиц в отдаленных и малонаселенных районах может быть значительно выше средней стоимости, поэтому доля отбираемых элементов будет ниже средней; напротив, в крупных городах, где наблюдается большее разнообразие условий занятости и экономической активности, эти доли могут быть выше. Под *отчетными областями* понимают части населения, для которых требуются отдельные оценки на основе обследования, обычно с определенной степенью точности в каждом случае. Эти требования могут диктовать различные доли отбора для разных отчетных областей (причем в небольших областях эти доли выше). Можно выделить три основных вида отчетных областей: географические области, например, городские и сельские районы или регионы страны (выборочная область обычно совпадает с этим видом или образует часть географической области); “рассеянные” области, например, отдельные подгруппы населения, сильно разбросанные по многим географическим областям; “смешанные” области, представляющие собой сравнительно сосредоточенные, но не вполне обособленные подгруппы населения. В географических и в меньшей степени в смешанных областях доли отбора вполне могут быть разными; применительно к рассеянным областям сделать это намного труднее, так как соответствующие подгруппы сильно разбросаны по географическим районам.

Существует ряд оснований предпочесть самовзвешенный отбор (во всей совокупности или по меньшей мере в каждой выборочной или отчетной области). 1) Взвешивание усложняет работу по обследованию: веса следует исчислить, сохранить в течение определенного периода, затем использовать при программировании и табулировании; об их существовании необходимо сообщить пользователям магнитных лент с данными; в опубликованных таблицах приходится показывать как взвешенные, так и невзвешенные объемы

выборки, если они существенно разнятся. 2) Беспорядочные веса, не связанные с дисперсией совокупности, ведут к увеличению дисперсии в результатах. Чем выше разброс весов, тем значительнее увеличение дисперсии. Такое увеличение дисперсии и сложности необходимо нейтрализовать на каждом этапе анализа. Поэтому чем продолжительнее и сложнее анализ, тем выше потери, связанные с отходом от самовзвешивания. 3) Взвешивание способно уменьшить гибкость и способность данной выборки удовлетворять различным целям и обслуживать разные обследования. Только отбор, близкий к самовзвешенному, может удовлетворять требованиям, вытекающим из различных проблем и разных обследований. 4) При самовзвешивании повторный отбор из одного списка не представляет труда, но когда предшествующий отбор базировался на неодинаковых вероятностях, последующий отбор усложняется. 5) Самовзвешенный отбор легче воспринимается пользователями — статистиками и широкой общественностью. 6) Следует отметить, что умеренное отступление от принципа самовзвешивания оказывает незначительное влияние на дисперсию. Сказанное означает, что отбор выше средней доли с целью оптимального распределения элементов, либо взвешивание для компенсации неполученных ответов или для повышения точности на этапе оценки оправданы лишь в том случае, если они связаны с относительно существенным отступлением от самовзвешивания.

Изложенные выше соображения в пользу самовзвешенного отбора (во всей совокупности или внутри каждой области) прежде всего относятся к широким обследованиям всего населения, например, обследованиям домашних хозяйств или рабочей силы. При обследовании экономических единиц, напротив, часто предпочтительнее выбирать более крупные предприятия с большей вероятностью.

Ступени отбора

При отборе элементов следует избегать лишних ступеней. Дополнительные ступени оправданы лишь в тех случаях, когда обеспечивается явное снижение затрат, уменьшение дисперсии и организационные преимущества.

При многоступенчатых схемах территориальные единицы часто отбирают с вероятностями, пропорциональными их размеру. На последней ступени конечные единицы можно отобрать с вероятностями, обратно пропорциональными размеру. Подобная схема имеет двойное преимущество, обеспечивая: а) самовзвешенный отбор конечных единиц и б) почти равномерную нагрузку внутри тер-

риторий. Нагрузка будет колебаться настолько, насколько были приблизительными показатели размера единиц при отборе выборки. Иногда для полного выравнивания нагрузки заранее фиксируют число подлежащих отбору конечных единиц, а не выборочные доли, обеспечивающие общий самовзвешенный отбор. Такая процедура способна привести к ряду сложностей и проблем, и ее применение практически никогда не бывает оправданным при обследованиях рабочей силы.

При обследованиях рабочей силы наиболее подходящими показателями размера обычно служат последние оценки численности населения. Где это возможно, следует также воспользоваться количеством домашних хозяйств или численностью населения старшего определенного возраста.

Можно также отметить, что на практике процесс отбора желательно *централизовать*, насколько это возможно. Отбор непосредственно участниками опроса часто ведет к серьезным искажениям.

5. ОТБОР ВО ВРЕМЕНИ

В парагр. 2 была подчеркнута важность обеспечения представительного характера выборки для обследования рабочей силы не только в пространственном, но и временном аспекте. В настоящем параграфе рассмотрены вопросы отбора элементов выборки в свете указанного требования. Целесообразно провести различие между двумя основными видами обследований рабочей силы (см. парагр. 3 гл. 10): а) фазовыми, или случайными, обследованиями, направленными на получение информации, представляющей долговременный интерес и относящейся к средним условиям или характеристикам, которые преобладают в течение некоторого периода времени; б) непрерывными обследованиями, направленными на получение текущих оценок уровней и тенденций.

В первом случае основное соображение сводится к следующему: для того чтобы данные, полученные применительно к определенному периоду в ходе обследования, можно было распространить на более продолжительный отрезок времени, сам этот период должен быть представительным с точки зрения более продолжительного отрезка времени. Точно так же период обследования сам должен быть достаточно длинным, чтобы охватить сезонные и иные колебания во времени: предпочтительнее раздвинуть временные рамки обследования, нежели проводить крупное обследование, ограниченное одним произвольно выбранным отрезком времени. Кроме того, для оценки, а также правильного сглаживания сезонных и иных колебаний период обследования необходимо разделить на

более мелкие отрезки (подсерии), и для каждого из них обследовать представительную выборку.

При продолжающемся обследовании цель заключается в получении текущих оценок (а также показателей динамики) через определенные периоды, например, раз в месяц, квартал или год. Поэтому обследование обычно организуют в виде продолжающейся серии раундов, и в ходе каждого из них получают отдельные оценки, относящиеся к установленному периоду. Как и в предшествующем случае, раунд обследования можно подразделить на несколько подраундов. В каком-то смысле разовое или случайное обследование схоже с отдельным раундом непрерывного обследования. Еще один вопрос, который требуется рассмотреть при организации такого обследования, — это степень независимости выборок для различных раундов, а также степень их корреляции и пересечения. Это зависит от “характера ротации”, принятой в обследовании.

Отбор элементов при разовых, или случайных обследованиях

Общее время и продолжительность обследования

Очевидно, что время проведения обследования в целом зависит от его задач. Если, к примеру, обследование призвано обеспечить статистические данные для разработки общенационального плана развития, то его следует завершить (включая стадии обработки и подготовки отчета) до начала разработки плана; однако не задолго до начала, с тем чтобы результаты переписи по возможности не устарели. Вообще говоря, такие обследования проводят для получения информации о структурных и других глубинных характеристиках, представляющих долговременный интерес. Предполагается, что результаты отражают средние условия в течение сравнительно продолжительного периода, границы которого можно определить лишь приблизительно. Поэтому при выборе времени проведения следует избегать нетипичных ситуаций.

Кроме того, период обследования должен быть растянут во времени настолько, чтобы были должным образом учтены и сглажены сезонные, циклические и случайные изменения преходящего характера. По названным причинам период обследования, как правило, охватывает целый год или по меньшей мере все основные сезоны года. Случаи, когда период обследования превышал бы год, нам неизвестны. Есть, однако, основания утверждать, что для того, чтобы период обследования отражал средние условия в течение всего периода, когда результаты обследования представляют интерес или

будут использованы, сам период обследования должен быть максимально возможной продолжительности и охватывать подвыборки, представительные в пространственном отношении в течение каждого отрезка времени. Такую схему называли “подвижной выборкой”, главная идея которой состоит в аккумулировании данных, собранных в течение длительного периода с целью уловить средние условия за данный период. Эффективное накопление данных предполагает, что подвыборки не пересекаются.

Деление периода обследования на подпериоды

В большинстве случаев весь период обследования целесообразно разбить на отрезки времени, которые можно назвать “подпериодами”. Этот вопрос рассмотрен в парагр. 3 гл. 10. В рамках каждого подпериода обследуется выборка, представительная в пространственном отношении, после чего можно получить отдельные оценки для подпериода. Каждый подпериод может охватывать отрезок времени равной продолжительности, например, месяц или квартал; или же подпериоды могут совпадать с основными сезонами, представляющими интерес, например, сельскохозяйственная страдная пора и межсезонье. Данные различных подпериодов можно агрегировать и получить средние результаты для всего периода обследования; путем сопоставления результатов подпериодов можно оценить сезонные колебания и другие тенденции. В качестве полезных иллюстраций приведем примеры из практики ряда стран.

В ходе 32-го раунда (1977 — 1978 г.) Национального выборочного обследования в Индии основное внимание было сосредоточено на занятости, безработице и потребительских расходах. Период обследования в один год был разбит на квартальные подпериоды. Выборка из первичных единиц (деревни и жилые кварталы) была распределена равными долями между четырьмя подпериодами. Сбор данных в каждом подпериоде был организован так, что каждый месяц квартала обследовали примерно одинаковое количество первичных единиц. Сбор данных проводили непрерывно в течение года.

Обследование рабочей силы в Шри-Ланке проходило в течение года с мая 1980 г. по апрель 1981 г. Оно охватило выборку примерно в 1000 переписных кварталов (первичных единиц отбора), в каждом из них для обследования в течение года были отобраны по 10 жилых единиц. Обследование проводили в четыре подраунда в течение года. В отличие от индийского обследования в ходе каждого подраунда обследовались жилые единицы во всех примерно 1000 первичных единиц. В каждой выборочной территории (блоке) в ходе

подпериодов охвату подлежало разное число жилых единиц; две жилые единицы на блок в первом и третьем подпериоде, и три — во втором и четвертом, т.е. всего по десять в каждой первичной единице в течение года.

И в Пакистане обследование рабочей силы, которое проводили в сентябре 1978 г. — августе 1979 г., было распределено на четыре подпериода. Выборки были отобраны в виде “взаимопроникающих” подвыборок равного объема и в каждый квартал обследовали одну подвыборку.

Распределение периода обследования на подпериоды обладает рядом преимуществ. Так как в каждом случае обследуется выборка, представительная в пространственном отношении, все обследование является представительным как в пространственном, так и в временном отношении в течение года обследования. Другое важное достоинство состоит в том, что сравнение результатов по подпериодам дает дополнительную информацию о сезонных изменениях и тенденциях в течение года, помимо наличия отдельных оценок по меньшей мере основных показателей для каждого подпериода. Имеются и организационные преимущества. К примеру, сбор данных можно проводить более равномерным темпом, его удастся лучше контролировать и регулировать; обработку данных и оформление результатов можно осуществлять последовательно по каждому подпериоду. Основной недостаток описываемой системы — высокие затраты, связанные с обследованием представительной и разбросанной выборки в каждом подпериоде. Дополнительные затраты ниже, если в каждом подпериоде обследуют лишь подмножество первичных или других единиц отбора на более высоких ступенях, как это делают в рамках охарактеризованного выше индийского обследования. Дополнительные затраты выше, когда в каждом подпериоде охватывают всю выборку первичных единиц (или более высоких ступеней отбора), как это имело место в Шри-Ланке. Удельные затраты (на единицу отбора) для получения общих оценок (например, за год) существенно возрастают, когда одно и то же множество конечных единиц обследуют неоднократно, например, в ходе каждого подпериода. Этой схемы придерживались при обследованиях рабочей силы, например, на Филиппинах и в Республике Корея (см. ниже).

Пересечение выборок для подпериода

В рамках одного обследования выборки для подпериодов могут соотноситься по-разному:

1. Один вариант — это независимая выборка для каждого подпериода. На практике это обычно означает наличие единого плана

отбора для всех подпериодов, который предусматривает, что выборки совершенно различны и не пересекаются ни на одной из ступеней. В каждом подпериоде обследуют разное множество первичных элементов отбора, а следовательно различны также все прочие элементы отбора, включая конечные элементы — домашние хозяйства, жилые единицы и т.п.

2. В каждом подпериоде можно обследовать разное множество конечных единиц, но все или некоторые элементы первичной и других ступеней отбора могут быть общими для подпериодов. Такие выборки не пересекаются с точки зрения конечного отбора, но они не являются независимыми.

3. В качестве модификации предыдущего варианта подмножество конечных единиц можно сохранять от одного периода к другому, а остальные замещать новыми. Это ротационная схема, и она связана с более тесной, по сравнению с предыдущим вариантом, корреляцией между выборками для различных подпериодов в зависимости от характера ротации/пересечения.

4. Наконец, можно обследовать ту же выборку конечных единиц в каждом подпериоде.

Выбор варианта зависит от целей обследования и соображений, связанных с затратами и эффективностью отбора. От обследования могут потребоваться четыре вида оценок:

оценки уровней для каждого подпериода в отдельности, например, месячные и квартальные оценки;

общие оценки для всего периода, агрегированные для разных подпериодов;

оценки частой разности или изменений от одного подпериода к другому, например, из месяца в месяц, квартальные или сезонные колебания;

оценки валовой разности или изменений на уровне отдельных индивидов или домашних хозяйств.

Если не считать выигрыша в точности, связанного с комбинированным и другими специальными методами оценки (см. парагр. 6), точность оценок уровней для отдельных подпериодов от пересечения выборок для них не зависит. Иначе обстоит дело с другими оценками. Оценки для всего периода обследования (агрегированные по подпериодам) наиболее эффективны, когда выборки для составляющих его подпериодов независимы или не пересекаются. Дело в том, что положительная ковариация связанных между собой или пересекающихся выборок ведет к увеличению дисперсии объединенного результата. Следовательно, вариант (1) более всего подходит для получения общих оценок, агрегированных по различным

подпериодам. При схемах (2) и (3), т.е. частично пересекающихся выборках, положительная ковариация взаимосвязанных выборок способствует повышению точности измерения сезонных колебаний. Это достоинство еще нагляднее при варианте (4); когда одна и та же выборка повторно обследуется в ходе каждого подпериода. Но при этом страдает точность, с которой можно агрегировать результаты по подпериодам при *данном общем числе опросов* в ходе обследования. При варианте (4) и в меньшей степени — (3) в связи с неоднократным обследованием возрастает также нагрузка на респондента.

Важно отметить, что для измерения сезонных и других колебаний на агрегированном уровне (т.е. для сравнения агрегированных оценок для разных подпериодов) *не обязательно* иметь пересекающиеся или взаимосвязанные выборки. Совершенно независимые выборки, извлеченные из той же совокупности, также дадут оценки изменений, но, как правило, менее точные. Чтобы учесть валовые изменения на *индивидуальном* уровне, например, сезонные закономерности экономической активности отдельных лиц, необходимо проследить за теми же индивидами на протяжении разных подпериодов. Точно так же, чтобы оценить динамику некоторых переменных на уровне индивидов за весь период обследования (например, статус экономической активности или доход за год, полученный сложением данных за четыре квартальных подпериода), необходимо либо: а) прибегнуть к более длительному отчетному периоду обследования, чем весь период (год), представляющий интерес, либо б) неоднократно обследовать тех же индивидов, когда используются более краткие отчетные периоды обследования. Первый вариант более экономичен, но его недостаток в том, что он дает информацию более низкого качества, так как предполагает необходимость вспоминать данные за продолжительный отчетный период, а это может быть сопряжено с проблемами.

Выбор характера ротации зависит от конкретных целей и относительной важности различных видов оценок. Во многих ситуациях главным требованием является наличие выборки достаточного объема, с тем чтобы результаты, полученные при агрегировании данных по подпериодам, были в полной мере детализированы в географическом отношении. В этом случае явно предпочтительнее вариант (1), использование полностью независимых или непересекающихся выборок для подпериодов (как в примере Индии, приведенном выше). Вариант (2) можно расценить как попытку шире распространить выборку конечных единиц — на элементы большего числа ступеней или всех ступеней — для того, чтобы получить более точные оценки на отдельных подпериодах (как в примере Шри-Лан-

ки). Вариант (3) представляет собой скорее компромиссное решение, удовлетворяющее различным целям. Издержки, связанные с вариантом (4), могут быть довольно значительными, и к нему следует прибегать лишь в том случае, когда главной целью обследования определенно является учет изменений на уровне индивидов. Примерами последней схемы служат обследования рабочей силы на Филиппинах в 1976/77 г. и обследования экономически активного населения в Республике Корея в 1977 г. в последующие годы. По меньшей мере в первом из упомянутых обследований, как сообщали, возникли проблемы, связанные с изменениями внутри выборки, элементы которой пытаются проследить в ходе всех подпериодов (Rao, 1985, p. 120).

Период обследования и отчетный период

Период обследования относится к отрезку времени, в течение которого собирают информацию, отчетный период — это время, к которому относится сама информация. Следует учесть, что под сказанным выше относительно “представительности во времени” имеются в виду, строго говоря, охваченные *отчетные периоды*, а не соответствующие периоды обследования. По названной причине важно четко уяснить, каким образом при обследовании охватывают отчетный период.

Рассмотрим обследование, при котором сбор данных осуществляется непрерывно и равномерно в течение года на основе скользящего отчетного периода продолжительности “ x ” (исчисляется в долях года). Очевидно, что информация, полученная от единиц, которые были учтены в самом начале обследования, относится к интервалу x , непосредственно предшествующему периоду обследования. Следовательно, весь период, к которому относятся данные обследования, это x плюс обследования в один год. Число единиц, по которым получена информация, в течение $(x + 1)$ лет неодинакова. Оно возрастает с нуля до максимума в первый x , остается максимальным в следующие $(1 - x)$ и, наконец, уменьшается до нуля в последний x . Почти ничего не меняется в том случае, когда x — это отчетный период, например, в одну неделю. Но ситуация резко меняется, когда x — это продолжительный отчетный период, равный, например, году. Информация в этом случае относится к двум годам: году до обследования и году обследования. Количество единиц, представляющих ретроспективную информацию, которая относится к любому моменту в течение этих двух лет, имеет треугольное распределение, пик которого приходится на начало периода обследо-

ния. Следовательно, обследование непосредственно не дает оценок для одного определенного года: это потребовало бы неэффективного инверсного взвешивания информации в зависимости от числа единиц, охваченных в каждый момент времени в течение года, и отказа от информации, относящейся к отрезку вне года. С другой стороны, невзвешенные результаты можно использовать непосредственно для изучения месячных или сезонных колебаний в том случае, когда безразлично, к какому из двух лет относятся данные. Охарактеризованные выше трудности не возникают, когда используют фиксированный отчетный период, т.е. один твердый календарный год для всех респондентов. Осложнения при фиксированном отчетном периоде возникают в связи с удлинением периода воспоминания для некоторых респондентов, о чем говорилось в парагр. 3 гл. 10.

Повторяющийся, или взаимопроникающий отбор

Один из вариантов плана выборки, который оказался пригодным для многих целей, — это взаимопроникающий, или повторяющийся отбор. Когда из одной совокупности на базе одинакового процесса отбора извлекают две или более выборки, их называют взаимопроникающими. Отбор не обязательно должен быть независимым, а уровни взаимопроникновения могут быть разными в соответствии со ступенями многоступенчатого отбора (Kendall and Buckland, 1980, p. 98). Например, если множества элементов первичного отбора извлекают на этой ступени, соответственно взаимопроникающими и будут эти выборки; возможно дальнейшее взаимопроникновение выборок и на второй ступени отбора и т.д. Схемы отбора, основанные на взаимопроникновении выборок, полезны во многих отношениях, например, для оценки дисперсии, оценки и контроля ошибок регистрации, сопоставления альтернативных процедур и т.д. Хороший обзор применений этой схемы содержится у Лахири (Lahiri, 1957). Выборки для различных подпериодов обследования можно запланировать в качестве взаимопроникающих, как отмечалось выше в примере, относящемся к Пакистану.

Обычно термины “повторяющиеся” и “взаимопроникающие” употребляют как синонимы. Но если быть более точным, то первый уже второго: к выборкам первого типа предъявляют дополнительное требование, а именно, их необходимо обследовать *одновременно*, в ходе того же обследования при сопоставимых условиях (Kendall and Buckland, 1980, p. 61 and p. 167). Общую выборку можно разделить на повторяющиеся подвыборки, каждая из которых представительна в пространственном отношении, и все они

обследуются одновременно в течение одного периода. Результаты различных повторов сравнивают и сопоставляют с целью получить представление о надежности данных и, при соответствующей схеме, оценить ошибки выборки и регистрации.

Дополнительные соображения при отборе для непрерывных обследований

Продолжающиеся обследования призваны давать оценки, относящиеся к нескольким отрезкам времени. Поэтому при таких обследованиях центральное место принадлежит понятию “раунда”. В зависимости от характера требуемых оценок и их относительной значимости возможна “ротация” выборки от одного раунда к другому. Возможные схемы ротации подробнее рассмотрены ниже.

Каждый раунд можно, разумеется, разделить на подраунды (подпериоды), о чем говорилось выше. Отмеченные выше соображения о пересечениях между различными подпериодами, а также о повторяющемся или взаимопроникающем отборе и т.п., относятся также и к раундам непрерывного обследования. Но следует отметить, что в этом случае цель разбивки на подпериоды заключается скорее в усилении операционного контроля и обеспечении равномерного распределения выборки во времени, нежели в получении отдельных оценок для подпериодов. Поэтому на практике результаты по подпериодам обычно агрегируют для получения оценок, относящихся ко всему периоду. Как указывалось выше, это удастся осуществить более эффективно при *непересекающихся* выборках для подпериодов. И действительно, чаще экономичнее разделить элементы первичного отбора, которые подлежат обследованию в раунде; на непересекающиеся подмножества, представительные в пространственном отношении, и ограничить каждый подпериод одним подмножеством.

Схемы ротации выборок

Ротация выборок означает, что от одного раунда к другому некоторые или все элементы выборки замещаются новыми элементами. В социальных обследованиях схема ротации обычно “симметрична”. Это означает, что новые множества элементов вводятся в выборку через регулярные интервалы, а после введения каждое множество сохраняют в выборке или выводят из нее по единой схеме. Как только такая система принята, она обеспечивает фиксированную степень пересечения выборок от одного раунда к другому

и между двумя раундами, разделенными данным отрезком времени. Несколько примеров помогут пояснить сказанное.

1. Начиная с 1977 г. выборка для обследования, проводимого два раза в год в Гонконге, базировалась на 10% -ной "главной выборке" на основе промежуточной переписи 1976 г. Она разделена на 15 равноценных подвыборок. При каждом полугодовом раунде обследуют две подвыборки: одну из числа обследованных в предшествующем раунде, а другую — новую. Сказанное означает, что между двумя последовательными раундами пересечение выборок равно 50% (на деле оно несколько меньше из-за неполученных ответов, перемещения единиц и других изменений); между двумя раундами, разделенными периодом в шесть месяцев или более, пересечение выборок отсутствует (Hong Kong, 1980).

2. Ежеквартальное обследование рабочей силы в Испании базируется на двухступенчатом отборе, причем на первой ступени элементами выступают секции, на второй — жилые единицы. Если не считать крупных пересмотров, секции, как правило, остаются неизменными от одного квартального раунда к другому. На каждом раунде обследуют 5/6 жилых единиц из предыдущего раунда и 1/6 жилых единиц (из той же секции). Следовательно, на уровне первичного отбора фактически существует полное пересечение выборок. На конечной ступени отбора выборки для двух последовательных раундов пересекаются на 5/6 общего числа единиц; для раундов, разделенных кварталом, пересечение равно 4/6; для раундов, разделенных двумя кварталами, оно составляет 3/6 (50%) и т.д.; для раундов, разделенных 15 месяцами и более, пересечение на уровне конечных единиц отсутствует (Spain, 1978). Чтобы охват периода раунда был равномерным, его выборка разделена на три независимые месячные подвыборки, каждая из которых представительна в отношении изучаемого населения определенной территории и охватывает 1/3 секций квартальной выборки. Кроме того, для организации контроля выборка дополнительно разбита на недельные порции секций, подлежащих обследованию в течение каждой недели.

3. Другие члены Европейского сообщества также проводят непрерывные обследования рабочей силы: ежеквартальные раунды в Нидерландах и Великобритании, как в Испании; ежегодные обследования (в весенний квартал) — в других странах. В большинстве обследований прибегают к постоянным схемам ротации, однако степень пересечения в них существенно различна. Например, пересечение выборок от одного года к другому равно 75% в ФРГ и Греции, 2/3 — в Дании и Франции, 50% — в Португалии, 1/2 — в

Италии (как в Испании), 25% — в Ирландии и Люксембурге (EUROSTAT, 1988). По мере увеличения интервала между раундами степень пересечения линейно уменьшается, за исключением Бельгии, где пересечение в 30% относится только к двум годам подряд, но не к последующим годам. Представляет интерес схема обследования в Великобритании. По сути дела, выборка здесь состоит из двух частей:

а) Квартальный компонент, охватываемый обследованием на постоянной основе и дающий представление о тенденциях внутри этого периода. Адрес остается в выборке на протяжении пяти квартальных раундов подряд, что означает высокую степень (80%) пересечения выборок от одного квартала к другому и позволяет более точно оценивать тенденции. Пересечение линейно уменьшается до 20% между двумя раундами, разделенными годом, и до нуля — впоследствии.

б) “Добавочный” компонент обследуется ежегодно в весенний квартал; его объем достаточно велик (он примерно в три раза больше, чем обычный квартальный компонент), чтобы можно было получить надежные базисные оценки *уровней* весной каждого года. В этой части выборки пересечение из года в год равно 1/3 объема.

4. Ротация в канадском обследовании рабочей силы также предполагает частичное замещение элементов выборки таким образом, что выборка для каждого месячного раунда представляет собой вероятностную выборку изучаемой совокупности. Ротация происходит на каждой ступени многоступенчатого отбора, принятого в обследовании. Конечная единица отбора — домашнее хозяйство — замещается каждые шесть месяцев, тогда как элементы на более высоких ступенях остаются в выборке более продолжительное время (Canada, 1976). Это означает, что от одного месячного раунда к следующему отбрасывают 1/6 часть домашних хозяйств и замещают в выборке новыми единицами, тогда как остальные 5/6 сохраняются из предшествующего раунда. На следующей ступени отбора смена из месяца в месяц затрагивает значительно меньшую долю элементов, причем эта доля последовательно убывает по мере перехода к более высоким ступеням отбора.

5. Примером из числа развивающихся стран может служить квартальное обследование рабочей силы в Индонезии (САКЕРНАС), которое было возобновлено в качестве отдельного обследования с 1986 г. Используемая схема ротации особенно проста: каждый квартал обследуют две подвыборки; в следующем квартале одну из них отбрасывают и обследуют новую. Иными словами, подвыборка подвергается обследованию два квартала кряду; между этими кварта-

лами пересечение равно 50%, затем оно отсутствует. Система лучше подходит для изучения квартальных, а не годовых изменений.

Приведенные примеры иллюстрируют относительно простые схемы ротации конечных единиц отбора. Выборка состоит из "n" подвыборок; в начале каждого периода обследования вводят новую подвыборку, а каждая подвыборка подвергается обследованию n периодов (раундов). Степень пересечения выборок между раундами линейно уменьшается по мере увеличения интервала между ними. Для двух выборок, разделенных "i" периодами, пересечение конечных единиц отбора равно $(n - i) / n$. После того как значение $i = n - 1$, пересечение равно нулю. В связи с неполученными ответами, перемещениями и другими измерениями, и особенно когда применяют концепцию охвата де-факто (наличного населения), пересечение на практике окажется ниже, чем это вытекает из формулы. Как правило, степень пересечения значительно выше для элементов на более высоких ступенях отбора, и она уменьшается по мере перехода к более низким ступеням. Во многих переписях фиксированную выборку конечных единиц сохраняют на протяжении сравнительно длительного времени. Часто раунды делят на подраунды, с тем чтобы обеспечить более равномерное распределение выборки во времени и улучшить контроль за операциями. Для этой цели выборку для раунда обычно разбивают на равные и непересекающиеся подвыборки, каждая из которых представлена в пространственном отношении (для определенной территории).

Для варьирования степени пересечения выборок и ее изменения по мере увеличения интервала между раундами прибегают к более сложным схемам ротации. Ниже приведены еще три примера.

6. При ежемесячном текущем обследовании населения США (United States, 1978) применяют так называемую схему "4—8—4". Это означает, что в начале каждого месяца к обследованию подключают новую подвыборку (именуемую "ротационной группой"); ее обследуют в течение четырех месяцев, затем исключают из выборки на восемь месяцев, затем обследуют еще четыре месяца, прежде чем окончательно с ней расстаться. Следовательно, за 16-месячный период конкретную единицу опрашивают в течение восьми месяцев; можно также увидеть, что на протяжении любого месяца обследуют восемь таких подвыборок — одну из них впервые, одну — во второй раз, и т.д. вплоть до группы, опрашиваемой в восьмой раз. Из месяца в месяц меняют две подвыборки, причем одну из них исключают временно, чтобы обследовать восемь месяцев спустя, а другую — окончательно. Иначе говоря, пересечение выборок от месяца к месяцу равно 75%. Пересечение сокращается до 50% через два меся-

ца, до 25% — через три месяца, затем оно равно нулю до того момента, когда оно снова возрастает при введении новых единиц.

Основная цель применения такой сложной схемы — добиться высокой степени пересечения выборок от года к году, с тем чтобы получить более точные оценки годовых изменений. При этой схеме пересечение становится ненулевым после девяти месяцев и остается таковым до 16-го месяца, составляя 50% для уровней, разделенных 12 месяцами.

Сказанное относится к конечным единицам (жилым единицам или их небольшим гнездам). На уровне “несамопредставительных” элементов первичного отбора ротация, по сути, отсутствует.

7. Обследование рабочей силы в Финляндии проводят квартальными раундами, каждый из которых разделен на месячные подпериоды. Каждый месяц привлекают новую подвыборку, которую обследуют пять раз в течение 16-месячного периода: в первый, четвертый, седьмой, 13-й и 16-й месяц после введения (Salmi and Kiiski, 1984). Легко увидеть, что каждый месяц обследуют пять таких подвыборок (что равно общему числу обследований любой подвыборки). В отличие от предыдущего примера пересечение выборки между подпериодами (из месяца в месяц) отсутствует: этим подразумевается, что учет помесечных колебаний (в отличие от квартальных) не является главной целью финского обследования. Однако, из одного квартала в следующий квартал пересечение равно 80%. Это значительно больше, чем 25%-ное квартальное пересечение в прежнем примере, что позволяет более точно оценивать квартальные колебания. Пересечение из года в год равно 40%, что несколько ниже, чем в США.

8. Последующий пример относится к обследованию рабочей силы в Японии. Нынешняя выборка базируется на переписи населения 1980 г., в качестве первой ступени отбора используют переписные участки, а в качестве второй (последней) — жилые единицы. Ротацию переписных участков производят по схеме “4—8—4”, описанной выше применительно к текущему обследованию населения в США. В каждом участке выборки отбирают два множества жилых единиц, и одно из множеств обследуют два месяца подряд, а на последующие два месяца заменяют другим множеством, если этот участок остается в выборке. Когда участок возвращается в выборку еще на четыре месяца в следующем году, схему повторяют (Rao, 1985, p. 83).

Хотя последняя схема и сходна с применяемой при текущем обследовании населения в США, она отличается от нее в двояком отношении. Во-первых, схема “4—8—4” относится здесь к ротации пер-

вичных элементов, тогда как в США эти элементы остаются неизменными, а ротация затрагивает конечные единицы. Это отражает то обстоятельство, что в США выборка базируется на небольшом числе элементов первичного отбора, каждый из которых представляет собой крупную инвестицию, тогда как в японском обследовании рабочей силы используют большое число (3000) относительно мелких первичных элементов. Второе различие состоит в том, что в японской выборке каждое домашнее хозяйство или каждого индивида обследуют лишь четыре раза за 16 месяцев, в отличие от восьми раз в США. Это сделано, чтобы уменьшить нагрузку на респондентов.

Некоторые факторы, определяющие выбор схемы ротации

Главное соображение связано со статистическими требованиями, определяемыми характером необходимых оценок и их сравнительной важностью. Когда основная задача состоит в агрегировании данных за некоторый период времени, лучше всего подходят непересекающиеся выборки. (Вот почему раунды обычно делят на непересекающиеся подпериоды, как это сделано в приведенных выше примерах по Испании и Финляндии.) При применении особых "составных" типов оценок (см. следующий раздел) можно повысить точность оценок и для пересекающихся выборок. Такие преимущества для пересекающихся групп конечных единиц могут иногда быть довольно значительными (Woodruff, 1963), но, как правило, они довольно скромны. Однако применительно к оценкам *изменений* от периода к периоду для пересекающихся выборок они бывают весьма существенными. Разумеется, валовые изменения можно учесть лишь в том случае, если обследованию из периода в период подвергаются те же элементарные единицы, как это имеет место при обследовании группами (панелями).

Степень пересечения, а следовательно, и схема ротации зависят также от продолжительности периода или периодов, для которых учитываются изменения. Более простые планы выборок (1) — (5), охарактеризованные выше, рассчитаны прежде всего на учет изменений между смежными периодами или раундами обследования, так как степень пересечения выборок для этих периодов максимальна, а затем убывает. Более сложные схемы (6) — (8) ведут к большему пересечению выборок для отстоящих друг от друга периодов, т.е. для учета изменений за год при текущем обследовании населения в США и японском обследовании рабочей силы или квартальных изменений, например, в Финляндии.

При выборе схемы ротации следует, разумеется, принимать во внимание стоимостные и организационные факторы. Замена выборки ведет к росту затрат. Сюда входят затраты на подготовку материалов для отбора, сам отбор, перемещение и замена интервьюеров, выявление новых респондентов и установление контакта с ними. Указанные затраты, как правило, выше, когда речь идет о замене элементов первичного отбора или элементов на других высоких ступенях отбора. По соображениям статистической эффективности и уменьшения нагрузки на респондентов целесообразнее заменять элементы более низких, чем более высоких ступеней отбора. По этим же причинам ротация, как правило, происходит быстрее на уровне конечных единиц и менее быстро либо вовсе отсутствует на уровне элементов первичного отбора.

Сохранение тех же конечных единиц в выборке также влечет за собой проблемы. Это ведет к увеличению нагрузки на респондентов, что в свою очередь может вызвать неполучение ответов и снижение качества данных. Ответы могут оказаться смещенными неизвестным образом. Примером может служить отмеченное во многих обследованиях рабочей силы со схемами ротации так называемое "смещение в связи с ротацией групп"; несмотря на все прочие факторы, результаты, как выясняется, зависят от того, сколько раз выборку опрашивали ранее. Установлено, что обусловленные смещениями различия наиболее отчетливо проявляются при сопоставлении подвыборок, опрошенных впервые, и подвыборки, подвергавшихся опросу ранее один или несколько раз (см. например, United States, 1978, p. 83—85). Мера целесообразных повторных опросов респондентов зависит от социо-культурных условий проведения обследования, его назначения и содержания, характера проводящей его организации и подготовки привлеченных интервьюеров. Во многих развивающихся странах, особенно в сельской местности, население нередко охотно сотрудничает с официальными или общественными организациями, проводящими статистические обследования. Однако такое сотрудничество нельзя считать гарантированным. В более развитых странах и городах сотрудничают, как правило, менее охотно. Проблемы могут также возникнуть в связи с колебаниями доли полученных ответов от раунда к раунду, даже когда речь идет об информации, относящейся к одним и тем же лицам. Это особенно заметно проявляется, когда разрешается получать ответы "по доверенности".

При сохранении той же выборки можно столкнуться и с другой проблемой, а именно, с трудностями при попытке проследить местонахождение и идентичность респондентов. В ряде случаев остроту

проблемы можно несколько ослабить путем охвата наличного населения (подход де-факто) и выбора более стабильных структурных элементов выборки (жилых единиц, подъездов и т.п.) вместо более мобильных семей и индивидов. Это, разумеется, уменьшит реальное пересечение элементов (индивидов и домашних хозяйств) между раундами.

Замещение выбывших единиц. Рассмотрим обследование, в котором от раунда к раунду происходит ротация (замещение) жилых единиц в рамках фиксированного множества гнезд элементов выборки. Наступит момент, когда все жилые единицы в небольших гнездах будут "использованы". Чтобы продолжить процесс в этой ситуации, можно либо: а) повторно использовать жилые единицы, которые уже находились в выборке, либо б) заменить использованные гнезда другими, ранее не обследованными. Первый вариант проще и дешевле, но иногда его считают нежелательным в связи с дополнительной нагрузкой на респондентов и возможным "смещением" ответов при неоднократном опросе одних и тех же респондентов. Второй вариант способен увеличить смещение самой выборки в результате постепенного исключения небольших элементов, если только не применены особые процедуры, призванные предупредить такого рода смещение.

Эти особые процедуры могут оказаться весьма сложными. Дело в том, что на практике во многих случаях прибегают к отбору гнезд или территорий на основе "переменной вероятности без возвращения", а этот метод, если его неоднократно применять к одной и той же совокупности, теоретически сложен. Необходимы более простые, практические методы. Один из них, пожалуй, наиболее простой, заключается в том, чтобы сгруппировать небольшие гнезда в достаточно крупные элементы еще до отбора, и использовать эти крупные элементы для отбора выборки и ротации. Указанный метод был применен в текущем обследовании населения в США (United States, 1978, p. 25). Охарактеризованную процедуру можно модифицировать, с тем чтобы уменьшить число выборочных гнезд, с которыми приходится оперировать в течение срока существования главной выборки.

Контроль за объемом выборки. Еще один специальный вопрос, который требует упоминания в связи с выборкой для непрерывных обследований, — это необходимость обеспечить, чтобы ее объем оставался со временем на требуемом уровне. Когда схема и фиксированные доли отбора остаются неизменными в течение длительного периода, наблюдается тенденция к автоматическому увеличению объема выборки в связи с естественным приростом численности

изучаемой совокупности в той мере, в какой это находит отражение в основе для отбора. Для того, чтобы объем выборки оставался постоянным, необходимо периодически корректировать доли отбора. Простейший метод — это уменьшить общие доли отбора (обычно на последней ступени) на основе соответствующего постоянного множителя и по мере отбора новой выборки для каждого раунда или группы раундов.

Метод панелей и система разделенной панели

Метод панелей означает использование фиксированной выборки элементарных единиц. Эта система позволяет учитывать такие валовые изменения, как перемещение индивидов между различными категориями экономической активности, различными профессиями и занятиями, разными отраслями или секторами экономики. Кроме того, путем наблюдения за одним и тем же множеством индивидов в течение длительного периода становится возможным подлинное изучение динамических аспектов перемен и причинно-следственных связей. Недостатки метода панелей таковы: высокие затраты; трудности при слежении за индивидами в течение длительного времени; убывание первоначальной выборки в связи с миграцией, смертью или иными изменениями; возможность “старения” панели и утраты ею представительности; увеличение нагрузки на респондентов; возможность увеличения числа ошибок из-за “смещения” ответов при неоднократном опросе. Несмотря на перечисленные проблемы, при определенных ситуациях и для известных целей замены продолжительному наблюдению не существует.

Сочетание схем ротации и панели можно назвать “системой разделенной панели” (этот термин предложил Лесли Киш). Идея заключается в том, чтобы привести панель фиксированных единиц в “классическую” схему симметричной ротации. Сама ротационная часть схемы может иметь и не иметь пересечения элементов. Система разделенной панели обладает рядом потенциальных преимуществ. Относительные объемы выборки в ее панельной и ротационной частях можно определять и корректировать более независимо и гибко. Одинаковую степень пересечения во всей выборке можно поддерживать на единообразной и постоянной основе на протяжении длительного периода. Постоянная панель дает возможность изучить динамические и общие изменения. В то же время путем сопоставления с переменной (ротационной) частью выборки можно получить представление о последствиях убывания и ухудшения па-

нели. Таким образом, система разделенной панели потенциально обладает специфическими преимуществами схемы ротации и схемы панели, одновременно позволяя избежать некоторых проблем.

Необходимо помнить, что сложные схемы с краткими раундами и подпериодами требуют жесткого контроля за сроками проведения работ и темпами сбора информации. На практике встречаются многочисленные обстоятельства, которые ведут к задержкам и утрате контроля над этими аспектами обследования. При отсутствии надежной связи и хороших транспортных средств определение места и момента возникновения проблем требует времени. Но даже когда проблемы выявлены, не всегда легко устранить их, например, когда персонал обследования не так-то легко переместить из одного района в другой по лингвистическим или этническим причинам и/или из-за значительных затрат и расстояний. Такого рода проблемы часто вызывают серьезные трудности при проведении обследований в развивающихся странах. Поэтому при выборе структуры обследования следует всегда стремиться к практичности и простоте.

6. ПРОЦЕДУРЫ ОЦЕНКИ

Всеобъемлющее непрерывное обследование рабочей силы можно использовать для получения широкого круга оценок для различных групп населения и с разной степенью агрегирования.

Во-первых, речь идет о различных *показателях*, например, степень вовлечения в рабочую силу, уровень безработицы, статус занятого, профессиональный состав и отраслевая принадлежность работающих и т.д. Показатели могут различаться по *требуемой* частоте и точности, а также *возможности* их достижения при данных масштабах и структуре обследования.

Во-вторых, оценки требуются обычно не только применительно ко всему обследуемому населению, но и к отдельным географическим и административным районам, городской и сельской местности, национальным образованиям, различным категориям населения, например, половозрастным и этническим группам, а внутри каждой категории необходима различная степень агрегирования. Эти категории и уровни именуют отчетными областями. Последним могут соответствовать и могут не соответствовать четко выделенные выборочные области в ходе планирования и отбора выборки (см. парагр. 4).

В-третьих, оценки могут потребоваться для различных типов *единиц анализа* — индивидов, домашних хозяйств, семей и общин.

В-четвертых, при выборочных обследованиях важно проводить различие между оценками долей, средних, уровней и других видов

коэффициентов, когда числитель и знаменатель получены при тех же условиях в рамках одного обследования; и оценками *сводных* величин, например, общего числа безработных, когда результаты обследования необходимо распространить на всю совокупность с помощью внутренней и внешней информации. Выборочные обследования скромных масштабов сами по себе часто способны обеспечить оценки сводных показателей для всего населения со значительно меньшей точностью (в смысле дисперсии и смещения) по сравнению с оценками долей, средних, уровней и аналогичных коэффициентов.

Наконец, виды оценок можно различить во временном аспекте. Как отмечалось выше, возможны оценки:

текущего уровня — для каждого отчетного периода или раунда обследования, например, месяца или квартала;

средних или агрегированных показателей для более длительных периодов — таких как год;

тенденцией или чистых изменений от одного периода к другому, например, из месяца в месяц, сезонных колебаний, изменений по годам;

валовых изменений на индивидуальном уровне, например, сдвигов между категориями экономической активности.

Следует также упомянуть долговременный учет динамики и причинно-следственных связей на основе наблюдения за фиксированным множеством индивидуальных единиц на протяжении некоторого периода времени.

Перечисленные выше виды оценок могут выступать в любом сочетании, что открывает огромные возможности. При выборе процедуры оценки в каждом случае следует учитывать предъявляемые к ней особые требования. В последующих разделах изложены основные принципы и некоторые важные аспекты процедур оценки при непрерывных обследованиях рабочей силы. Они включают также несколько более узкие требования, предъявляемые к разовым, или случайным обследованиям.

Простые несмещенные оценки

Очень важно иметь возможность на основе данных обследования получать простые несмещенные оценки, даже если эти последние можно впоследствии улучшать или модифицировать при отыскании окончательных оценок. В практике обследований термин “простые несмещенные оценки” употребляют в приблизительном значении. При этом, по сути, подразумевают следующее: а) оценки получают

непосредственно на основе результатов обследования, не прибегая к данным, внешним по отношению к нему, взвешивая каждое наблюдение в обратной пропорции к вероятности его отбора для выборки; б) полученные таким путем оценки являются приблизительно несмещенными в статистическом смысле, по меньшей мере, при среднем и большом объеме выборки. Сказанное не обязательно означает, что фактически полученные оценки свободны от искажений, связанных с реализацией выборки и ошибками в ответах. Понимается лишь, что если бы та же процедура была применена при тех же условиях для повторного обследования, то при отсутствии таких ошибок ожидаемое среднее значение результатов было бы приблизительно равно значению для совокупности, представляющей интерес.

Такие оценки возможны только при вероятностном отборе, т.е. при использовании выборок, полученных таким образом, что любой элемент совокупности имеет известную ненулевую вероятность быть отобранным. Необходимо также, чтобы проблемы реализации выборки, например, получение ответов и недостаточный охват, не искажали эти вероятности, а их влияние на вероятности можно было количественно оценить с высокой степенью уверенности. Хорошие простые оценки предполагают также, что любые коррективы, которые придется впоследствии вносить для их уточнения, не окажутся большими.

Способность получать хорошие "простые несмещенные оценки" означает, что обследование было запланировано и проведено должным с научной точки зрения образом — отсюда важность таких оценок.

Плановые веса

Процедура оценки на этом этапе сравнительно проста: каждое наблюдение взвешивают на величины, обратно пропорциональные вероятности его отбора для выборки. Эти веса называют "плановыми". Полученные взвешенные данные можно агрегировать любым образом. Как отмечено в парагр. 4, обследования рабочей силы часто планируют в виде самовзвешенных. В этом случае выборочные оценки можно получить без взвешивания данных, а результаты затем распространить на всю совокупность с помощью постоянного множителя, обратно пропорционального общей доле выборки. Эта доля может колебаться от одного отбора или одной области к другой, но выборка в рамках каждой области должна быть самовзвешенной. В этом случае для каждой области можно получить невзвешенные оценки, а затем рассчитать взвешенную должным образом сумму

для всей совокупности (с помощью весов, обратно пропорциональных доле отбора в данной области). В общем же случае более удобно присвоить соответствующий вес каждой *отдельной единице* анализа, внесенной в компьютер, даже когда к отдельным группам единиц применимы одинаковые веса. Это облегчает и делает более гибкими компьютерную обработку и табулирование данных.

Корректировка в связи с неполученными ответами

В определенных случаях плановые веса необходимо скорректировать в связи с особыми проблемами, возникающими в процессе реализации плана. Наиболее распространенной и важной является корректировка из-за неполученных ответов. Неполучение ответов отражается на результатах обследования различным образом:

1. В той мере, в какой не давшие ответа единицы систематически отличаются от тех, кто дал ответ, распределение в выборке искажается, и результаты оказываются смешанными.

2. Оценки для всей совокупности, непосредственно полученные на основе выборки, окажутся смещенными вниз, если не будут приняты во внимание неучтенные единицы.

3. Неполучение ответов фактически означает уменьшение объема выборки, а следовательно, увеличение дисперсии выборки.

4. Помимо "общего" неполучения ответов, когда от некоторых единиц выборки информации вообще не поступает, возможно также "частичное" неполучение в виде неполной информации от некоторых единиц. Эта проблема обесценивает и усложняет анализ данных обследования.

Никакая корректировка не способна полностью устранить последствия неполучения ответов. Цель различных корректировок — по возможности уменьшить связанные с ним искажения. Обычно прибегают к корректировкам трех видов: замена не давших ответа единиц другими единицами примерно с теми же характеристиками; добавление отсутствующих значений на основе частичной информации, имеющейся в отношении этих или других единиц; корректировка плановых весов с целью уточнения, насколько это возможно, распределения и общего объема выборки.

Замена единиц — это распространенная практика во многих обследованиях, но часто она приносит мало пользы и способна повлечь за собой утрату контроля над операциями по сбору данных. Поэтому мы в общем рекомендуем избегать ее в большинстве случаев. Добавление отсутствующих значений часто необходимо и целесообразно при комплексных обследованиях, но следует предостеречь против

огольного и широкого использования и этой процедуры. (Общая характеристика такого рода процедур содержится в работе ООН, 1982, с. 99—102). Мы рассмотрим здесь методы корректировки, упомянутые последними, а именно, поправки на неполученные ответы, основанные на оценке и предполагающие изменение плановых весов.

Основная процедура корректировки плановых весов в связи с неполучением ответов такова. Выборки разбивают на “ячейки”, относительно однородные с точки зрения некоего важного критерия, применимого к единицам, которые дали ответ и не дали его. Для каждой ячейки вычисляют соответствующий коэффициент ответов (исчисление коэффициентов ответов для взвешенных и невзвешенных выборок рассмотрено в гл. 14). Оценку для ячейки или веса для каждой индивидуальной единицы затем распространяют на совокупность с помощью обратной величины коэффициента ответов для каждой ячейки.

Важное значение имеет правильный выбор ячеек и категорий, используемых для корректировки. Существенно, разумеется, чтобы между этими категориями можно было распределить как единицы, давшие ответ, так и не давшие. Сами эти категории должны быть содержательными и важными, с тем чтобы их распределение в изучаемой совокупности нашло правильное отражение в выборке (взвешенной). Они также должны быть по возможности внутренне однородными. Целесообразно использовать категории, которые, вероятнее всего, имеют разные коэффициенты ответов и различаются по ключевым показателям. Они должны быть достаточно большими по размеру, чтобы применяемые коррективы в весах не были слишком разными и большими. Дело в том, что взвешивание, как правило, ведет к увеличению дисперсии, хотя и способствует уменьшению смещения. Наконец, процедура должна быть простой.

В качестве категорий, применяемых для корректировки результатов из-за неполучения ответов, часто используют географическое деление, крупные районы или группы территориальных единиц. Этим достигается то, что распределение полученной выборки соответствует распределению отобранной выборки, по меньшей мере на достаточно высоком уровне агрегирования. Иногда корректировку производили на уровне небольших конечных территориальных единиц; однако, такое использование очень небольших районов для корректировки выборочных весов, как правило, не рекомендуется из-за связанных с этим больших и переменных весов. Взвешивание возможно также на базе нетерриториальных подклассов или групп в зависимости от характеристик домашних хозяйств и индивидов и

при наличии необходимой для группировки информации как для давших ответы единиц, так и для тех, кто их не дал.

Следующая проблема относится к определению коэффициентов ответов. Для самовзвешенной выборки коэффициент ответов это попросту число успешно опрошенных единиц, деленное на общее число отобранных единиц (сумма опрошенных и неопрошенных единиц). На практике при точном подсчете могут возникнуть трудности. Например, при гнездовой выборке число отобранных элементарных единиц может быть твердо не установленным и неизвестным; точно так же выборочные списки часто содержат "пропуски", которые не всегда легко отличить от подлинных случаев неполучения ответов. (Эти проблемы подробнее рассмотрены в парагр. 3 гл. 14 и в работе Марквардта, 1984.) При несамовзвешенных выборках количество единиц, используемое для исчисления коэффициента ответов, следует взвешивать на величину, обратно пропорциональную вероятностям отбора единиц, давших ответ и не давших ответа. Все веса опрошенных единиц затем увеличивают на величину, обратную этому взвешенному коэффициенту ответов.

Использование внешних весов

Исчисление итоговых величин

Путем описанного выше умножения всех выборочных данных на величину, обратную вероятностям отбора единиц, можно, конечно, на базе обследования рассчитать соотношения, а также агрегированные показатели для совокупности. Этот метод вполне подходит для исчисления пропорций, средних коэффициентов и других соотношений в рамках выборочных обследований, но он часто крайне неудовлетворителен при оценке общих показателей для совокупности, особенно при малых и умеренных объемах выборки. Дело в том, что при многоступенчатом отборе объем выборки — это случайная переменная, а агрегированные величины, исчисленные непосредственно на базе выборки, могут включать значительные ошибки выборки. Проблема может еще более осложниться, когда оценки относятся к подклассам совокупности, разбросанным по выборочным гнездам так, что их количество в выборке при ее планировании контролировать трудно. Кроме того, оценки агрегированных показателей искажаются в результате ошибок в охвате и других аналогичных ошибок.

Правильный подход поэтому во многих случаях состоит в том, чтобы исчислить соотношения непосредственно на базе выборки (с

возможными поправками, описанными ниже), а агрегированные показатели для совокупности на основе этих соотношений рассчитывать, прибегая к дополнительной надежной информации, внешней по отношению к самому обследованию. Приведенный ниже пример призван помочь уяснить эту идею.

Предположим, требуется оценить общую численность экономически активного населения (скажем, Y). Предположим, что в *выборке* количество экономически активных лиц равно y , а число лиц в соответствующей базовой совокупности равно x . (Допустим, что выборка самовзвешенная, а общий множитель — F , т.е. норма отбора — $1/F$.) Наконец, предположим, что благодаря надежному внешнему источнику нам известно, что общее базовое население равно X . Тогда агрегированный показатель Y можно исчислить двояким путем:

простая несмещенная оценка — $Y = Fy$

оценка на основе соотношения — $Y = (y/x)X$.

Вторая оценка, будучи технически смещенной, может иметь меньшую дисперсию и среднюю квадратическую ошибку. Ценность и применимость этой процедуры зависят, однако, от нескольких факторов. Во-первых, коэффициент корреляции между y и x должен быть положительным и предпочтительно большим, скажем, не менее 0,6 или 0,7. Во-вторых, X должно быть известно со значительно большей степенью точности, нежели (Fy); иными словами, внешний источник должен быть надежным. В-третьих, X в совокупности и x в выборке должны базироваться, по сути, на идентичном охвате и учете. Существенное различие между ними может привести к значительному смещению в оценке для всей совокупности. Это часто предполагает, что значение x также следует брать из внешнего источника, а не из данных, учтенных непосредственно при обследовании (Kish, 1965, section 6.5D).

Контроль данных о населении

Такой контроль относится к внешним последним оценкам данных, которые можно использовать для того, чтобы различные коэффициенты и соотношения, полученные в ходе обследования, распространить на всю совокупность единиц. Применительно к обследованиям рабочей силы контроль необходим в отношении данных не только о всем населении, но и важных подгрупп, например, половозрастных, этнических и т.п. социальных категорий, сельского и городского населения, жителей отдельных регионов. Такие данные обычно получают с помощью краткосрочных прогнозов,

исходя из самой последней переписи населения, дополненной информацией из административных источников и обследований, характеризующих динамику различных аспектов народонаселения. Стандартные методы оценивания и прогнозирования численности населения подробно описаны в различных источниках, и здесь воспроизводить их нет необходимости. В частности, можно сослаться на серию ООН "Руководства по методам прогнозирования численности населения" (Руководство I относится ко всему населению, Руководство III — прогнозированию половозрастного состава, Руководство VIII — прогнозированию в разрезе городской и сельской местности). Полезное обсуждение проблемы контроля данных о населении на ином уровне содержится в работе, выпущенной в США (United States, 1966).

Оценка соотношений с помощью внешних весов

Использование внешних данных о численности населения с целью распространения выборочных соотношений на всю совокупность — это частный случай общей методики использования внешних весов для корректировки выборочных результатов. Цель при этом заключается в использовании внешних весов для повышения точности простых несмещенных оценок, а возможно, и для внесения поправок на смещения, обусловленные ошибками охвата и другими погрешностями при реализации выборки.

Основная процедура такова. Выборку разбивают на несколько частей с учетом некоторых важных характеристик элементов. Распределение выборки по этим характеристикам сопоставляют с распределением всей совокупности по этим же характеристикам. С учетом соотношений этих распределений получают корректирующие множители, или веса, которые, будучи применены к выборочным показателям, приводят выборочное распределение в соответствие с внешними стандартами. Такое корректирующее взвешивание можно применить в виде единой операции или же последовательно для различных характеристик и групп характеристик либо даже итеративно для обеспечения наиболее полного согласования.

В сущности, одинаковая процедура применима и к оценкам соотношений, и к оценкам совокупных агрегированных величин. Единственное различие состоит в том, что в первом случае соответствующие внешние данные имеют форму относительных *распределений*, тогда как во втором речь идет об абсолютных величинах совокупности. Во многих ситуациях относительные значения

можно получить с большей точностью, чем абсолютные величины. Но в общих случаях *относительные* веса отдельных элементов выборки корректируют одинаковым образом.

Полезно привести несколько примеров. При обследовании рабочей силы в Таиланде уровни и соотношения, исчисленные на базе выборки, взвешивают согласно общим данным о населении по пятилетним возрастным группам (21 группа) отдельно по виду территории (три группы: муниципальные районы, "санитарные районы" и сельские районы) и пяти регионам. Следовательно, процедура связана с большим числом (более 300) контрольных ячеек и реализуется в виде единой операции (Thailand, 1985, p. 13—15).

При текущем обследовании населения в США прибегают к ряду корректировок. Во-первых, в рамках крупных ячеек, определяемых расовой принадлежностью и местом жительства, *учтенное переписью* население в единицах первичного отбора каждой ячейки сравнивают с общим населением данной ячейки. Это обеспечивает множитель, который следует использовать для корректировки, обусловленной тем, что обследуется только выборка из единиц первичного отбора. Цель при этом состоит в том, чтобы уменьшить вклад первой ступени отбора в дисперсию выборки. На втором шаге корректируют общенациональные выборочные оценки по ряду половозрастных и расовых групп населения, используя при этом независимые текущие показатели численности населения в разрезе всех этих групп. Третий шаг предполагает дальнейшую корректировку весов с помощью отдельных множителей, исчисленных для половозрастных (с пятилетней разбивкой) и расовых групп и отдельно — для каждой из восьми "ротационных групп" (т.е. подвыборок в зависимости от месяца их включения в выборку), что в общей сложности дает более 500 оцениваемых ячеек (United States, 1978, p. 58—61).

В принципе, внешние корректирующие веса можно применить к любой выборке, независимо от вероятностной природы и представительности самой выборки. Они способны скорректировать грубые ошибки и искажения в выборке. Однако они не гарантируют отсутствия смещения, если выборка в принципе плохого качества или же внешние данные сомнительной точности. Правильный подход состоит в том, чтобы запланировать и реализовать выборку должным образом, используя строгие вероятностные методы, а к корректирующим весам прибегать лишь тогда, когда известно, что внешний источник явно превосходит выборку с точки зрения информации о численности населения и его распределении между различными оцениваемыми ячейками.

Корректирующие факторы, значительно отклоняющиеся от 1,0, заставляют предположить неудовлетворительное планирование и реализацию выборки.

В статистически менее развитых странах надежные и оперативные данные часто отсутствуют, но особенно ненадежно их распределение по отдельным группам. По названной причине необходимо проявлять осторожность при применении внешних весов для корректирования выборочных результатов, и, конечно, избегать корректирования на слишком низком уровне классификации. Если точность внешних данных сомнительна, они скорее всего не подойдут для уточнения правильных в своей основе оценок, полученных при обследовании: их использование может не только не устранить искажения, но даже увеличить их число. Прибегать к внешним весам в таких случаях следует лишь для исправления серьезных недостатков выборки на достаточно высоком уровне агрегирования.

К сожалению, непродуманные и сложные корректировки выборочных результатов на базе внешних данных неудовлетворительного качества практиковали при обследованиях рабочей силы в ряде стран. Против такой практики следует настоятельно предостеречь.

Составные оценки при пересекающихся выборках

Результаты, полученные в разные периоды (раунды), при пересекающихся выборках характеризуются положительной корреляцией. Степень корреляции зависит от характера показателя и изучаемой совокупности, степени пересечения, ступени отбора, на которой имеет место пересечение, и, разумеется, от временного интервала между рассматриваемыми периодами. Применительно к многим показателям, представляющим интерес при обследованиях рабочей силы, можно ожидать достаточно высокую степень корреляции, особенно в краткосрочном аспекте. Цель процедур составной оценки в том, чтобы воспользоваться такой положительной корреляцией во времени между одинаковыми и смежными единицами для повышения точности оценок текущих уровней, и особенно оценок изменений показателей. Когда приходится агрегировать показатели за различные периоды (раунды), наличие положительной корреляции между смежными оценками, как правило, снижает точность агрегированных результатов. Составные оценки иногда можно использовать с целью отчасти компенсировать указанный недостаток.

Основная идея составных оценок состоит в замене единственной оценки взвешенной суммой оценок той же величины, полученных с помощью разных процедур или из различных источников.

Рассмотрим, к примеру, простую ротационную выборку, в которой некая случайная доля элементов переходит из одного раунда в другой, а остальная часть заменяется. Каждая из двух частей дает оценку текущего уровня с одинаковой точностью (если отвлечься от обычного влияния объема выборки), а при объединении обеих частей получают оценку, которая не зависит от степени пересечения. Однако при оценке *изменений* от одного раунда к другому пересекающаяся часть дает благодаря положительной корреляции сравнительно более точные результаты по сравнению с непересекающейся частью. Следовательно, оценку изменений можно улучшить, если придать первой части больший вес, чем второй. Киш (Kish, 1965, section 12.4) показал, что минимальную дисперсию разностей получают, уменьшив относительный вес непересекающейся части на множитель $(1 - R)$, где R — коэффициент корреляции между тем же множеством наблюдений за два периода. Обратное справедливо при оценке агрегированных величин или сумм за два периода. В этом случае точность можно повысить, если умножить относительный вес непересекающейся части выборки на $(1 + R)$.

Оценки текущих уровней можно также улучшить при пересекающихся выборках путем использования информации, полученной в ходе предыдущих раундов, и факта положительной корреляции между результатами двух периодов. Теоретически аспекты метода, применяемого при месячных обследованиях в США, описаны Вудраффом (Woodruff, 1963). В сущности, процедура предполагает получение улучшенной текущей оценки как взвешенной суммы двух составляющих: а) оценки, полученной непосредственно на базе текущего раунда; б) оценки для предшествующего раунда, скорректированной с учетом выявленного изменения между предшествующим и текущим раундами общей панели единиц, обследованных в обоих раундах. Это изменение можно выразить в виде разности между значениями, полученными в обоих раундах (например, United, 1978, p. 64), или их соотношения (Woodruff, 1963). Существуют методы оптимизации выбора весов с целью минимизации дисперсии составной оценки. Веса будут зависеть от степени соответствующей корреляции между методами. Чем больше коэффициент корреляции, тем больше относительный вес, придаваемый составляющей оценки (б).

При обследованиях рабочей силы некоторые показатели (например безработица), как правило, показывают менее тесную корреляцию между периодами по сравнению с другими показателями (статус занятости работающих). Это означает, что в идеале существует потребность в различных наборах весов, применяемых в со-

ставных оценках, для разных показателей. Но в большинстве случаев такое усложнение неоправданно (из-за того, что результаты сравнительно нечувствительны к точности применяемых весов), во всяком случае на практике это трудно осуществимо. Поэтому для всех показателей, представляющих интерес, можно воспользоваться одним набором компромиссных весов.

Некоторые специальные вопросы

Синтетические оценки

Учитывая важность информации об экономической активности населения, все более возрастает спрос на оценки, которые были бы не только оперативными, но и детальными, т.е. применительно к небольшим территориям и группам населения. Во многих странах наблюдается стремление к постоянному увеличению объема выборок при обследованиях. Существуют, однако, рамки, в пределах которых можно, а фактически и *следует*, держать объемы выборки из-за отрицательного влияния крупного объема выборки на затраты, качество и оперативность выдаваемых данных. Поэтому многие выборочные обследования обеспечивают текущие оценки с ограниченной степенью географической и иной детализации; в то же время данные можно аккумулировать во времени, с тем чтобы повысить степень детализации, но в ущерб оперативности.

Цель методов синтетических оценок заключается в получении одновременно детальных и оперативных оценок путем совместного использования информации из различных источников, включая переписи, административные органы и выборочные обследования. Теория и практика использования этих процедур быстро развивались в последние годы в связи с возросшим спросом на информацию, созданием более полных банков данных и совершенствованием компьютерной техники. (Анализ последних достижений см. в работе Платена и др., 1987; более ранний, но детальный пример их использования для оценок безработицы и жилищных условий в США содержится в работе Гонзалеса и Хоза, 1978; анализ процедур, в частности в связи с канадским обследованием рабочей силы, опубликован Статистическим бюро Канады.)

Необходимо отметить, что разработка хороших процедур синтетических оценок зависит от качества и количества имеющихся данных. Выбор методов оценок должен базироваться на эмпирических исследованиях и сопоставлении полученных результатов с имеющимися внешними стандартами. Однако эти методы становятся все

более необходимыми и полезными, и их разработку следует также поощрять в развивающихся странах.

Сезонные поправки к динамическим рядам

Статистика рабочей силы, как и многие другие данные, отражает регулярно повторяющиеся сезонные колебания, которые можно оценить на основе прошлого опыта. Исключив ту часть изменений, которую можно приписать обычным сезонным колебаниям, становится возможным наблюдать циклические и другие несезонные изменения в динамическом ряду. Доводы, обычно выдвигаемые в пользу исключения сезонных колебаний, состоят в том, что это позволяет: а) облегчить краткосрочное прогнозирование; б) соотнести динамические ряды с другими данными, внешними событиями и мерами экономической политики; в) добиться сопоставимости показателей рядов во времени; г) упростить данные, с тем чтобы их могли легче понять и истолковать пользователи, не имеющие солидной статистической подготовки.

Для исключения сезонной составляющей колебаний разработаны различные методы и модели. Наиболее известный метод, применяемый в ряде стран, это метод "X—11" Бюро переписей США (Shiskin, Young and Musgrave, 1965).

Совместимость ошибок

Наконец, кратко упомянем важный вопрос, который относится не только к данным о рабочей силе, но и к оценке и публикации результатов статистических исследований в любой области, особенно если речь идет о публикациях национального органа статистики. Это вопрос *совместимости*: данные, опубликованные на основе одного источника, должны быть совместимы с данными из другого источника, по меньшей мере когда можно явно рассчитывать на такую совместимость. Совместимость результатов предполагает, конечно, совместимость понятий, определений, классификаций и охвата в различных источниках. Ее можно улучшить, обеспечив совместимость структуры и методологии, а также повсеместным соблюдением определенных стандартов в отношении приемлемой и поддающейся измерению точности. А в некоторых случаях совместимость необходимо *делать обязательной*, особенно когда данные из различных источников собирает и публикует один официальный орган. Это означает, что данные должны соответствовать некоторым контрольным итогам (по меньшей мере на достаточно высоком уровне агрегирования), даже если при этом придется прибегнуть к несколь-

ко произвольному методу. Это еще один пример использования внешних весов на стадии оценки. Следует помнить, что статистические данные приносят пользу только когда их используют, если же пользователь сталкивается с разными цифрами для одного показателя, приведенными одним и тем же органом, это может вызвать только замешательство и раздражение.

Один из аспектов совместимости — это *аддитивность* оценок, полученных на различных уровнях агрегирования. Сказанное означает, что сумма показателей для взаимоисключающих областей должна быть равна оценкам для всей совокупности. В этом смысле оценки соотношений, будучи более точными, чем простые несмещенные оценки, могут не быть аддитивными; может оказаться желательным несколько скорректировать веса, применяемые на стадии оценки, с тем чтобы обеспечить аддитивность, во всяком случае на определенном уровне агрегирования и выше. Конкретная и полезная рекомендация содержится в работе Сантера (Sunter, 1975).

Здесь не представляется возможным детально обсудить различные вопросы, относящиеся к совместимости публикуемых официальных данных и упомянутых в предыдущих абзацах. Два момента относительно чисто механических аспектов совместимости в таблицах будут, однако, рассмотрены. Они едва ли когда-нибудь имеют большое значение по существу, но их следует иметь в виду при публикации и подаче данных, рассчитанных на пользователей-нестатистиков.

Во-первых, в отчетах об обследованиях рабочей силы часто содержатся таблицы, где, помимо прочих, фигурируют три вида данных: а) домашние хозяйства; б) главы домашних хозяйств; в) отдельные лица. Обычно расчетное число домашних хозяйств, полученное на базе внешнего источника, позволяет проконтролировать итоги (а), а расчетное число лиц, сгруппированных по полу, возрасту и т.п., — проконтролировать итоги (б) и (в). А так как каждое домашнее хозяйство должно иметь одну и только одну главу, оценки для (а) и (б) должны быть идентичны. Есть основания полагать, что в совместимом внешнем источнике, вроде расчетных данных переписи, дело будет обстоять именно так. Но если расчеты производились отдельно на базе внешних итоговых данных о домашних хозяйствах и населении, (а) и (б) могут точно не совпадать. Это связано с тем, что средний размер домашнего хозяйства, полученный в ходе обследования, может не совпадать с размером во внешнем источнике. Указанное затруднение можно разрешить путем *дальнейшей группировки* расчетного числа лиц из внешнего источника на основе того, является ли индивид главой домашнего хозяйства или нет, еще

до использования этого числа при расчете соответствующего показателя для всей совокупности на базе выборочного соотношения.

Во-вторых, оценки на базе выборки часто предполагают применение дробных (в отличие от целочисленных) весов. Из-за ошибок в округлении это может привести к небольшим, но раздражающим расхождениям в итогах одних и тех же показателей в разных таблицах. В ряде стран указанную проблему решают следующим образом. Рассмотрим группу единиц с одинаковыми дробными весами. Группу разбивают случайным образом на две части такой величины, чтобы при присвоении одной из них целочисленного веса чуть меньше дробного, а другой — целочисленного веса чуть больше дробного среднее значение обеих групп было равно действительному (дробному) весу. (Очевидно, что такая процедура предполагает, что дробный вес пересчитан, чтобы он превышал 1,0, так что ни одному из элементов выборки не присваивают нулевой вес.)

7. ИЗМЕНЕНИЕ ПЛАНА ВЫБОРКИ

Общие аспекты изменения плана выборки вместе с некоторыми примерами рассмотрены в парагр. 5 гл. 10. В настоящем параграфе обращено внимание на некоторые специфические стороны этого вопроса.

Использование иной основы выборки

Наиболее подходящий момент для серьезного изменения плана выборки возникает, когда доступны результаты новой переписи населения. Результаты переписи можно использовать для обновления выборки с учетом прироста населения, изменений в его распределении и характеристиках, изменений в размерах, границах и характере домашних хозяйств, пересмотра определений отчетных областей и т.д. Одновременно можно скорректировать охват и масштабы обследования. На базе такой измененной основы и смежных материалов выборки можно обновить полностью или частично.

Изменение объема и структуры выборки

Помимо изменения основы выборки пересмотр плана отбора предполагает анализ объема, распределения и стратификации выборки, частоты сбора данных, характера ротации элементов и других аспектов структуры выборки с учетом информации об издержках и дисперсии, а также организационных и содержательных требований. Целесообразно попытаться выявить и исправить *основные не-*

увязки, которые могут присутствовать в плане. Нередко даже приблизительная информация о дисперсии, издержках и запаздывании в получении результатов может предсказать, требуются ли серьезные коррективы в объеме и распределении выборки, а также частоте сбора информации. Уточнение схемы ротации элементов обычно требует более детальной информации и углубленного анализа.

Другой вопрос, который часто возникает при изменении плана выборки, — это коррективы, необходимые для получения более точных оценок для небольших и/или пересмотренных отчетных областей. К примеру, во многих обследованиях постепенно переходят от получения общенациональных оценок к оценкам на уровне регионов или провинций. Кроме изменения объема и распределения выборки, это может быть сопряжено с переосмыслением слоев и элементов отбора, с тем чтобы они соответствовали новому плану и новым отчетным областям. Аналогичные соображения возникают при изменении классификации местностей в зависимости от вида населенных пунктов, например, в связи с урбанизацией и расширением городских границ.

Усовершенствованные процедуры оценки

Улучшение процедур оценивания зависит от наличия более совершенной вспомогательной информации, которой можно воспользоваться при оценке, и информации относительно дисперсии и смещения оценок при ранее применявшихся процедурах. В отличие от других аспектов пересмотра плана выборки, процедуры оценки не влияют на организационную сторону обследования, а потому их можно спокойно корректировать на любых ступенях. Единственным ограничивающим фактором является необходимость обеспечить преемственность и сопоставимость результатов, а также тот факт, что, вообще говоря, процедуры оценки поддаются усовершенствованию лишь на базе серьезных исследований и анализа. Усовершенствование этих процедур затрагивает ряд моментов, например: разработку и применение улучшенных методов контроля данных о населении; уточненные веса для составных оценок на основе более надежной информации о корреляции во времени; разработку процедур для различных типов единиц, например, оценок применительно к индивидам, в отличие от оценок применительно к семьям или домашним хозяйствам; разработку или усовершенствование синтетических оценок для небольших территорий и групп населения.

Преемственность и изменение выборки

Изменение плана выборки не обязательно означает полную замену или новый отбор любой существующей выборки. В тех случаях, когда подготовка материалов для отбора и создания инфраструктуры сбора данных в намеченных районах потребовали значительных затрат, может оказаться весьма желательным не допустить новых затрат, несмотря на пересмотр выборки. Разработаны несколько методов, призванных обеспечить сохранение в выборке максимально возможного количества элементов старой выборки без потери ею своего объективно вероятностного характера (Kish and Scott, 1971). По аналогичным причинам иногда считают более предпочтительным снова дополнить существующую выборку, с тем чтобы она удовлетворяла изменившимся требованиям к отчетности, хотя полный пересмотр был бы более эффективен, если начать заново. Примером может служить расширение текущего обследования населения в США для получения более надежных оценок по штатам; в этом случае существующая выборка была дополнена новыми выборками с помощью сложных “зависимых” процедур, т.е. отбор новых элементов зависел от элементов, которые уже присутствовали в прежней общенациональной выборке.

Еще один аспект преемственности — это постепенное включение новой схемы ротации и новых элементов с целью избежать внезапных изменений в динамических рядах, получаемых в ходе обследования. Это предполагает составление графика, на базе которого старые территории, которые покидают выборку в соответствии со схемой ротации, заменяют новыми на основе сбалансированного подхода.

Более широкие аспекты изменения выборки

Изменение плана выборки может, разумеется, потребовать пересмотра более широких аспектов плана обследования и в свою очередь зависеть от них. Упомянем некоторые из этих аспектов.

Изменение выборки может быть обусловлено коренным пересмотром основного метода сбора данных. Это может быть связано, например, с переходом от почтового опроса к телефонному или личному опросу, изменением состава интервьюеров или организацией сбора данных (например, вместо стационарных счетчиков, работающих по одному, вводятся передвижные бригады интервьюеров).

При крупномасштабных обследованиях рабочей силы важным вопросом может быть их расширение с целью создания базы для проведения других смежных обследований. Это может сделать необхо-

димым принятие соответствующего компромиссного плана выборки, которая удовлетворяла бы более разнообразным требованиям.

Наконец, следует упомянуть, что серьезный пересмотр выборки может быть серьезным поводом для введения новой главной выборки, которая, как отмечено в парагр. 3, способна обслуживать потребности различных раундов обследования рабочей силы, равно как и других видов обследований.

Литература

- Canada. 1976. *Methodology of the Canadian Labour Force Survey*. Ottawa, Statistics Canada. Catalogue No. 71—526 Occasional.
- Casterline, J. (ed.). 1985. *The collection and analysis of community data*. The Hague, International Statistical Institute.
- Cochran, W. G. 1953. *Sampling techniques*. New York, Wiley.
- Deming, W. E. 1960. *Sample design in business research*. New York, Wiley.
- European Communities, Statistical Office (EUROSTAT). 1988. *Labour force survey, Methods and definitions*. Theme 3, Series E (Catalogue CA-49-87-874-EN-C) & Luxembourg, Office for Official Publications of the European Communities.
- Gonzalez, M.; Hoza, C. 1978. "Small area estimation with application to unemployment and housing estimates", in *Journal of American Statistical Association*, 73, pp. 7—15.
- Hansen, M.; Hurwitz, W.; Madow, W. 1953. *Sample survey methods and theory*. New York, Wiley.
- Hong Kong. 1980. *Report on March 1980 labour force survey*. Hong Kong. Census and Statistics Department.
- International Labour Office (ILO). 1986. *Statistical sources and methods*, Vol. 3. Economically active population, employment, unemployment and hours of work (household surveys). Geneva.
- Kendall, M. G.; Buckland, W. R. 1980. *A dictionary of statistical terms*. (Fourth ed.). London, Longman.
- Kish, L. 1965. *Survey sampling*. New York, Wiley.
- 1965a. "Sampling organisations and groups of unequal sizes", in *American Sociological Review*, 30 (4), pp. 564—572.

- 1977. "Robustness in survey sampling", in *Bulletin of the International Statistical Institute*, 47 (3), pp. 515—528.
- ; Hess, I. 1958. "On non-coverage of sample dwellings", in *Journal of the American Statistical Association*, 53, pp. 509—524.
- ; Scott, A. 1971. "Retaining unit after changing strata and probabilities", in *Journal of the American Statistical Association*, 66, pp. 461—470.
- Lahiri, D. B. 1957. "Observations on the use of interpenetrating samples in India", in *Bulletin of the International Statistical Institute*, 36 (3), pp. 144—152.
- Marckwardt, A. M. 1984. *Response rates, callbacks and coverage: The World Fertility Survey experience*. WFS Scientific Reports, No. 55, The Hague, International Statistical Institute.
- Murthy, M. N. 1967. *Sampling theory and methods*. New Delhi, Statistical Publishing Society.
- ; Roy, A. S. 1970. "A problem of integration of surveys: A case study", in *Journal of the American Statistical Association*, 65, pp. 123—135.
- 1975. "Development of the sample design of the Indian national sample survey during its first twenty-five rounds", in *Sankhya, the Indian Journal of Statistics*, 37, (C) 1, pp. 1—42.
- Platek, R. et al (ed.). 1987. *Small area statistics*. New York, Wiley.
- Rao, M. V. S. 1985. *Labour force and family living surveys in Asia and the Pacific*. Bangkok, Economic and Social Commission for Asia and the Pacific (ESCAP), Statistics Division, St/EsCAP/312.
- Salmi, H.; Kiiski, S. 1984. "Developing the quality of the labour force survey: The Finnish approach", in *Statistisk Tidskrift* (Stockholm, Statistics Sweden), 1984.1, pp. 16—26.
- Shiskin, J.; Young, A.; Musgrave, J. 1965. *The X-11 variant of the census method II seasonal adjustment program*. Technical paper No. 15. Washington, DC, US Bureau of the Census.
- Spain. 1978. *Labour force survey technical report: Design of the survey and quality evaluation of the data*. Madrid, Instituto Nacional de Estadística.
- Sukhatme, P. V. 1954. *Sampling theory of survey with applications*. Amex, Iowa College Press.
- Sunter, A. 1975. "A survey design strategy for a national statistical agency" in *Bulletin of the International Statistical Institute*, 46 (1), pp. 221—235.
- Thailand. 1985. *Report on the Labour Force Survey*. Bangkok, National Statistical Office.

- United Nations. 1960. *A short manual on sampling*. No. 9, Rev. 1, New York, UN Statistical Office.
- 1980. *Principles and Recommendations for population and housing censuses*. Series M, No. 67, New York, UN Statistical Office.
 - 1982. *Survey data processing*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.
 - 1984. *Handbook of household surveys*. Series F, No. 31, New York, UN Statistical Office.
 - 1986. *Sampling frames and sample designs for integrated household survey programmes*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.
 - 1986. *Manuals on methods of estimating population*: Manual of estimating total population for current dates (ST/SOA/Series. A/10); Manual III, Methods for population projection by sex and age (ST/SOA/Series.A/25); Manual VIII, Methods of projections of urban and rural population (ST/ESA/SER.A/55). New York, UN Statistical Office.
- United States. 1966. Atlantida Unit III — Population data controls, Series ISPO 1, No. 1-D. Washington, DC, US Bureau of the Census.
- 1978. *The current population survey: Design and methodology*. Technical Paper, NO. 40. Washington, DC, US Bureau of the Census.
- Verma, V. 1977. "Sample designs for the World Fertility Survey", in *Bulletin of the International Statistical Institute*, 47 (3), pp. 101—117.
- ; Scott, C.; O'Muircheartaigh, C. 1980. "Sample designs and sampling errors for the World Fertility Survey", in *Journal of the Royal Statistical Society, A*, 143 (4), pp. 431—473.
- Woodruff, R. S. 1963. "The use of rotating samples in the Census Bureau's monthly surveys", in *Journal of the American Statistical Association*, 58, pp. 454—467.
- World Fertility Survey. 1977. *Guidelines for country report no. 1*. WFS Basic Documentation, No. 8, The Hague, International Statistical Institute.
- Yates, F. 1949. *Sampling methods for censuses and surveys*. London, Griffin.

Глава 12. РАЗРАБОТКА И ПОСТРОЕНИЕ АНКЕТЫ

1. ВВЕДЕНИЕ

Важность работы по составлению анкеты

Один из наиболее важных аспектов разработки анкеты — точное определение подлежащих сбору данных и преобразование требований к ним и смежных понятий в конкретные вопросы, которые респондент может понять и на которые может правильно ответить. Разработка анкеты затрагивает все аспекты качества данных, включая полезность, своевременность и точность полученных результатов; она может также серьезно повлиять на издержки, связанные с обследованием.

Один из главных принципов совершенствования плана обследования заключается в перераспределении и концентрации ресурсов на тех направлениях, где они более всего могут повысить качество данных. Указанный принцип подчеркивает важность улучшения анкеты обследования, ибо во многих случаях с помощью незначительных затрат на эти цели можно добиться существенного уменьшения числа ошибок. Тем не менее в практике обследований очень часто разработке хорошей анкеты не придают должного значения и не выделяют для этого необходимых средств. Имеются многочисленные примеры обследований, когда в результате неудовлетворительного построения анкеты собирали ненужные или недостаточно качественные данные, происходили задержки в их обработке. Многие организации и лица, причастные к обследованиям, полностью не осознают сложность задачи разработки хорошей анкеты. Выразить сложные понятия и определения в виде ряда относительно простых вопросов далеко не просто. И на то имеется несколько причин. В обследовании, рассчитанном на широкий охват, одни и те же вопросы должны относиться к самым разнообразным ситуациям и различным респондентам. Задача доведения вопросов до респондентов и сбора требуемой информации децентрализована. Ее необходимо поручать интервьюерам, расположенным в различных местностях вдали от бдительного ока контролирующего органа, и это обстоятельство трудно компенсировать любым контролем. Интервьюеры по-разному относятся к делу и неодинаково понимают содержание и методы обследования. Языковые различия и варианты еще более усложняют задачу. Кроме того, анкета должна отвечать целям не только сбора точных данных, но и эффективной обработки и анализа результатов.

Поэтому для разработки анкеты необходимы различные навыки и квалификация: знание сути обследования; исчерпывающее знакомство с практическими условиями, в которых будет собираться информация; знание принципов, инструментов и методов построения анкеты, включая чертежные навыки и чувство формы; и особенно способность уделять внимание кажущимся мелочам. Такие навыки нечасто встречаются в организациях, проводящих обследования, и во всяком случае их эффективное использование предполагает должную организацию работы по составлению анкеты.

Несмотря на всю сложность этой работы, по-видимому, широко распространено убеждение, что анкету можно "составить" быстро и просто и что любой, кто обладает хотя бы некоторым "здоровым смыслом" и в целом знаком с практикой обследований, способен сделать это или во всяком случае может участвовать в такой работе. Подобная точка зрения, очевидно, связана с тем, что разработка анкеты, как представляется, не требует такой же степени специализации, как, скажем, планирование выборки, обработка данных и статистический анализ. На деле это не так, а распространенное убеждение о простоте задачи построения в принципе ошибочно и способно отрицательно повлиять на качество всего обследования.

Разумеется, масштабы и сложность задачи, связанной с разработкой анкеты, меняются от одного обследования к другому в зависимости от характера подлежащей сбору информации, способа сбора информации, вида респондентов и задействованных интервьюеров, предшествующего опыта и знаний. Так, в странах со сложившейся системой обследования рабочей силы задача может заключаться главным образом в периодическом улучшении и уточнении ранее использованной анкеты. В этом случае важная задача будет состоять в том, чтобы при всех вносимых изменениях сохранялась максимальная преемственность с прежними раундами, а разрыв в динамических рядах был бы минимален. Обычно это предполагает внесение в имеющиеся анкеты только существенных изменений на основе четко осознанных целей. Задача усложняется, когда в стране организуют новое обследование рабочей силы, но и в этом случае, по-видимому, существует возможность опереться на более или менее испытанные и сложившиеся методы и подходы. Дело в том, что обычно крупномасштабные национальные обследования принимают лишь на основе апробированных процедур, когда можно опереться на международные рекомендации, образцы материалов из разных источников, а также иностранный и национальный опыт проведения аналогичных работ. В этом случае разработка анкеты будет заключаться в основном в анализе имеющегося опыта; воз-

можно более точном определении целей обследования и требуемых результатов; разработке анкеты, удовлетворяющей конкретным требованиям; проверке и оценке проекта анкеты на небольшой, но достаточно представительной выборке респондентов, обычно на основе более или менее проверенных методологий сбора данных. Иногда при новых обследованиях, особых требованиях к данным и недостаточном прежнем опыте, на который можно опереться, может оказаться необходимым собрать от потенциальных респондентов дополнительную информацию и провести более тщательную оценку того, что является целесообразным в условиях, при которых осуществляют обследования. Могут потребоваться особые “качественные” методы опроса, например, неформальный “разговорный” опрос, целенаправленная беседа для выяснения отдельных вопросов, полуструктурированный опрос, с возможностью многих вариантов ответа, ключевых или специально отобранных респондентов. Полезными могут также оказаться по-разному организованные групповые опросы и изучение на основе наблюдения. Применение таких специальных методов все еще ограничено, но они получают все большее распространение. Эти специальные методы нами не обсуждаются, но можно сослаться на опубликованную работу Кесли и Кумара (1988 г., гл. 2—4).

Организационные предпосылки

Разработка анкеты обследования — это пошаговый процесс. Он начинается с первоначального определения общих целей обследования, далее уточняют и конкретизируют содержание обследования, которое преобразуют в конкретные вопросы, сведенные в одну или несколько анкет, а затем проводят работу на местах, обрабатывают, оценивают и анализируют данные. Этот процесс приобретает все более технический характер, и требуются соответствующие организационные и консультационные меры с учетом меняющихся требований. Хотя детали конкретных мероприятий меняются, существуют три главные организационные предпосылки надлежащей организации разработки анкеты. Особенно важно уделять внимание этим предпосылкам при крупных статистических работах типа общенациональных обследований рабочей силы.

Первое требование — это обеспечение надлежащих условий для *консультаций между пользователями и поставщиками данных*. Определение того, какую и когда собирать информацию — это, в общем, прерогатива пользователей; организаторы обследования должны исходить из этих требований. Тем не менее взаимодействие

сторон должно отвечать их обоюдным интересам. Знания организаторов обследования и специалистов могут помочь пользователям более точно и детально сформулировать их требования. В этом процессе цели обследования конкретизируются и могут даже претерпеть изменения по мере их анализа. Важно четко разграничить функции пользователей и организаторов обследования. Пользователям следует определить, *какая* нужна информация, т.е. ожидаемый результат. Решать, *как* собирать требуемую информацию (с точки зрения методологии и организации обследования, формулировки вопросов и т.д.) следует технически лучше подготовленным специалистам. Разумеется, необходимые консультации сторон следует проводить на всех этапах, но они должны в основном касаться содержательных результатов обследования, а также соответствующих затрат и ресурсов. Разработка анкет обследования в рамках “комитетов пользователей и разработчиков” может закончиться выдачей неудачного продукта, так как эти рамки не способствуют тщательному рассмотрению всех возникающих практических и технических аспектов.

Как только определены общее содержание и ожидаемые результаты обследования, разработка анкеты обследования приобретает все более конкретный и технический характер. Она требует участия различных специалистов, в том числе знающих суть обследования, условия сбора и обработки данных и, прежде всего, — принципы построения анкеты. Чтобы решить эту задачу, необходимо сформировать *бригаду разработчиков анкеты*, обладающих необходимой квалификацией. Таково второе требование.

Третье главное требование — это создать необходимые условия для *анализа, оценки и пересмотра* анкет. Анализ следует проводить на широкой основе с привлечением не только самих разработчиков проекта анкеты, но и пользователей соответствующего уровня, внешних специалистов, организаторов обследования и контрольного персонала. Анализ должен носить организованный и систематический характер. Следует, однако, отметить, что такая широкая группа — как правило, не самый подходящий орган для внесения в проект анкеты окончательных поправок или дополнений в конкретной форме. Дело в том, что пересмотр анкеты может быть весьма сложной задачей, требующей тщательного анализа многочисленных деталей. Такую задачу лучше поручить технически грамотной небольшой группе (подобно упомянутой выше бригаде по разработке анкеты), которая может затем доложить о результатах работы более представительной группе.

Содержание настоящей главы

Материал настоящей главы расположен следующим образом. В парагр. 2 и 3 рассмотрены содержательные аспекты разработки анкеты для обследований рабочей силы. Как и в других главах, предполагается, что обследование проводится на основе устного опроса домашних хозяйств и индивидов. В парагр. 2 дается общий обзор построения анкеты как пошагового процесса, который начинается с выявления и уточнения необходимой информации и ожидаемых результатов, включает конкретизацию аспектов обследования и заканчивается формулированием вопросов анкеты. В парагр. 3 приведены некоторые примеры логических схем анкеты из национальных обследований рабочей силы, призванные проиллюстрировать проблемы, возникающие при определении содержания анкеты.

В парагр. 4 и 5 рассмотрены технические стороны разработки анкеты. Эти принципы и методы в меньшей степени зависят от конкретного содержания обследования. В основе выбора структуры анкеты лежит принцип желательной степени стандартизации процесса опроса. В этой связи в парагр. 4 обсуждается формулировка вопросов и структуризация ответов. В парагр. 5 рассмотрены проблемы применения анкет при крупномасштабных обследованиях.

Всестороннее и детальное обсуждение принципов разработки и построения анкеты с многочисленными примерами содержится в работе ООН (1985 г.), представляющей собой одно из технических пособий, разработанных в рамках Программы ООН по обследованиям домашних хозяйств. Эта глава во многом опирается на содержащийся там материал. Работа, выполненная в связи с Обследованием рождаемости в мире, также содержит весьма полезный иллюстративный материал, особенно по вопросу построения и структуры анкет для устного опроса (World Fertility Survey, 1975) и разработки инструкций для опроса, контроля, учебы, редактирования, кодирования и обработки данных; эти материалы опубликованы в серии "Основная документация по обследованию рождаемости в мире", выпущенной Международным статистическим институтом. Полезное обсуждение указанных принципов содержится у Джабина (Jabine, 1983).

2. УТОЧНЕНИЕ СОДЕРЖАНИЯ И РЕЗУЛЬТАТОВ ОБСЛЕДОВАНИЯ

Основные соображения

Разработка анкеты начинается с детального и тщательного уточнения подлежащей сбору информации. Определяя содержание обследования, всегда желательно начинать с ожидаемого *конечного*

продукта, а затем идти в обратном направлении, чтобы выйти на инструменты и процедуры, необходимые для его получения. Следует также получить точные ответы на многие более широкие вопросы планирования и методологии обследования, имеющие отношение к условиям сбора и обработки данных. Они были обсуждены в гл. 10 и относятся к решениям, например, о целевой группе населения, подлежащей охвату, виде и частоте собираемой информации, возможной взаимосвязи обследования рабочей силы с другими обследованиями (как в содержательном, так и организационном плане), способе получения информации, организации и процедурах обработки данных. Содержание, форма и стиль используемой анкеты зависят от того, кто будет источником информации и как ее будут получать. Совершенно разные соображения лежат в основе анкеты, предназначенной, например, для устного опроса, в отличие от письменного или телефонного опроса. Точно так же существуют различия между обследованиями, при которых респондент отвечает только за себя или же может давать информацию и о других лицах; а также между обследованиями, которые проводят высококвалифицированные, подготовленные интервьюеры, и обследованиями, где заняты счетчики-непрофессионалы.

До детального определения показателей и вопросов обследования следует ясно и подробно уточнить определения понятий, четко и подробно установить, что включается в них и исключается из них и как эти определения будут применяться на практике. Общий недостаток многих обследований заключается в том, что в то время как в документах и отчетах детально определены и классифицированы сложные и широкие понятия, в анкетах и инструкциях для интервьюеров достаточно подробно не разъяснены основные и наводящие вопросы, необходимые для учета таких показателей. В этом случае теоретические упражнения, связанные с определением понятий, имеют мало смысла.

Еще одна существенная проблема — возможное применение при обследовании более чем одной анкеты. В рамках одного обследования рабочей силы может собираться информация о различных видах и уровнях единиц — индивидах, домашних хозяйствах, а в ряде случаев — и об округе. Кроме того, по некоторым проблемам информацию можно собирать для всей выборки, а по другим проблемам может быть достаточно данных по подвыборке. Частота сбора информации по разным проблемам также может быть неодинаковой. В зависимости от проблем могут также различаться виды респондентов, предоставляющих данные. По указанным и аналогичным причинам часто необходимо (и удобно) использовать в одном обследовании более одной анкеты.

Хотя подготовку подробного плана выпуска таблиц с полным описанием построения каждой таблицы можно отложить до окончания разработки схемы анкеты и кодирования, основное содержание планов выпуска таблиц и их анализа следует определить как можно раньше: в идеальном случае — одновременно с уточнением содержания обследования, и уж, конечно, до детальной проработки анкеты (United Nations, 1985, p. 35).

Схема выпуска таблиц должна содержать как минимум, наименование таблиц, показатели, которые в ней найдут отражение и будут использованы для классификации; группы населения, к которым будут относиться различные таблицы. Желательно также как можно подробнее показать классификационные признаки, хотя впоследствии здесь могут потребоваться уточнения, когда будет лучше известно выборочное распределение ответов на вопросы. На известной стадии разработки анкеты становится возможным подготовить макеты таблиц с указанием наименований, столбцов, строк и т.д. Такие макеты полезны, поскольку они позволяют выявить пробелы в подлежащей сбору информации. Они дают возможность также выявить ненужную информацию или излишние детали, включенные в проект вопросника.

Изложенное выше сохраняет силу несмотря на то, что в связи с совершенствованием обработки данных значение предварительного плана выпуска таблиц отчасти уменьшается, во всяком случае в развитых странах. Проводимые теперь обследования все более рассчитаны на гибкий анализ данных, а не только на выпуск определенного числа заранее намеченных таблиц.

Уточнение показателей обследования и логические схемы анкет

Показатели обследования

После определения характера подлежащей сбору информации и связанных с этим понятий, определений и группировок переходят к составлению подробного перечня и описания показателей обследования. Эта работа может проходить параллельно с выяснением ожидаемых результатов обследования в виде комбинированных таблиц и аналитических материалов. Как правило, при разработке анкет предпочтительнее выявить показатели обследования до формирования последовательности вопросов.

Полезно различать показатели обследования и пункты анкеты. Показатели характеризуют содержание получаемой в ходе обследования информации на микроуровне, тогда как пункты анкеты определяют точную форму, в которой информация подлежит сбору. Показатель может не соответствовать единственному пункту анкет

ты. Например, в интересах получения более точной информации в анкете может быть предусмотрена большая степень детализации пунктов, нежели это необходимо для конечного анализа. Возможны ситуации, когда по практическим соображениям, связанным со сложностью и размером анкеты, более детальный перечень показателей приходится свести к сокращенной (и, возможно, менее точной) серии вопросов, опустив некоторые детали или перенести их в инструкцию для интервьюеров или программу их обучения. И действительно, в зависимости от сложности и разнообразия ситуации с занятостью, подлежащей изучению, большинство анкет обследования рабочей силы нуждается в таком упрощении.

Следует упомянуть, что сами показатели обследования можно определить более и менее подробно, и в первом случае различие между показателями и фактическими вопросами сводится лишь к форме или формулировке. К примеру, показатель обследования “безработица” можно дезагрегировать на ряд показателей типа “ищущих работу”, “готовых приступить к работе” и т.д.

Тем не менее полезно помнить о различии между показателями, представляющими интерес, и конкретными вопросами, требующимися для их выявления. Желательно сперва составить полный перечень показателей, а преобразовать их в конкретную анкету лишь на более поздней стадии, после того, как перечень обсужден и утвержден. В отличие от перечня показателей обследования, при детализации анкеты следует учитывать массу других факторов, включая формулировки, размещение, порядок и другие аспекты формы, помимо основного содержания подлежащей сбору информации. Для ясности и удобства может, например, потребоваться параллельная серия вопросов для различных категорий респондентов, даже если информация, получаемая с помощью различных серий, одинакова по содержанию. Все эти детали способны затемнить суть дела и затруднить пользователям и интервьюерам контроль и оценку содержания обследования. Напротив, четкий перечень показателей с их определением и характеристикой того, что включается и что исключается, указанием совокупности, к которой они относятся, и возможных ответов и т.д. более точно и кратко передает содержание обследования. Такой перечень облегчает и разработку плана выпуска таблиц и других результатов обследования.

Логические схемы анкеты

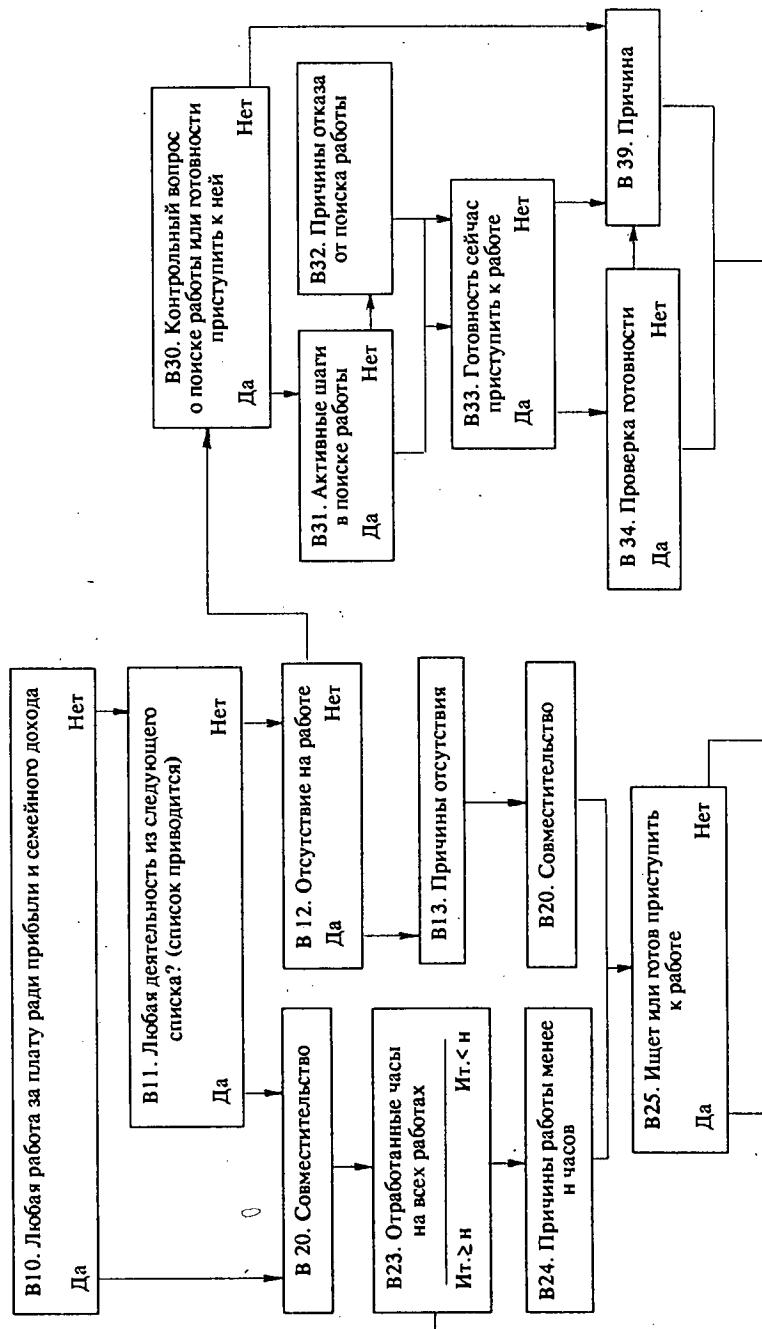
В различных частях логических схем, приведенных в части I настоящего руководства, по сути, приведены достаточно подробные перечни показателей обследования и элементов информации, под-

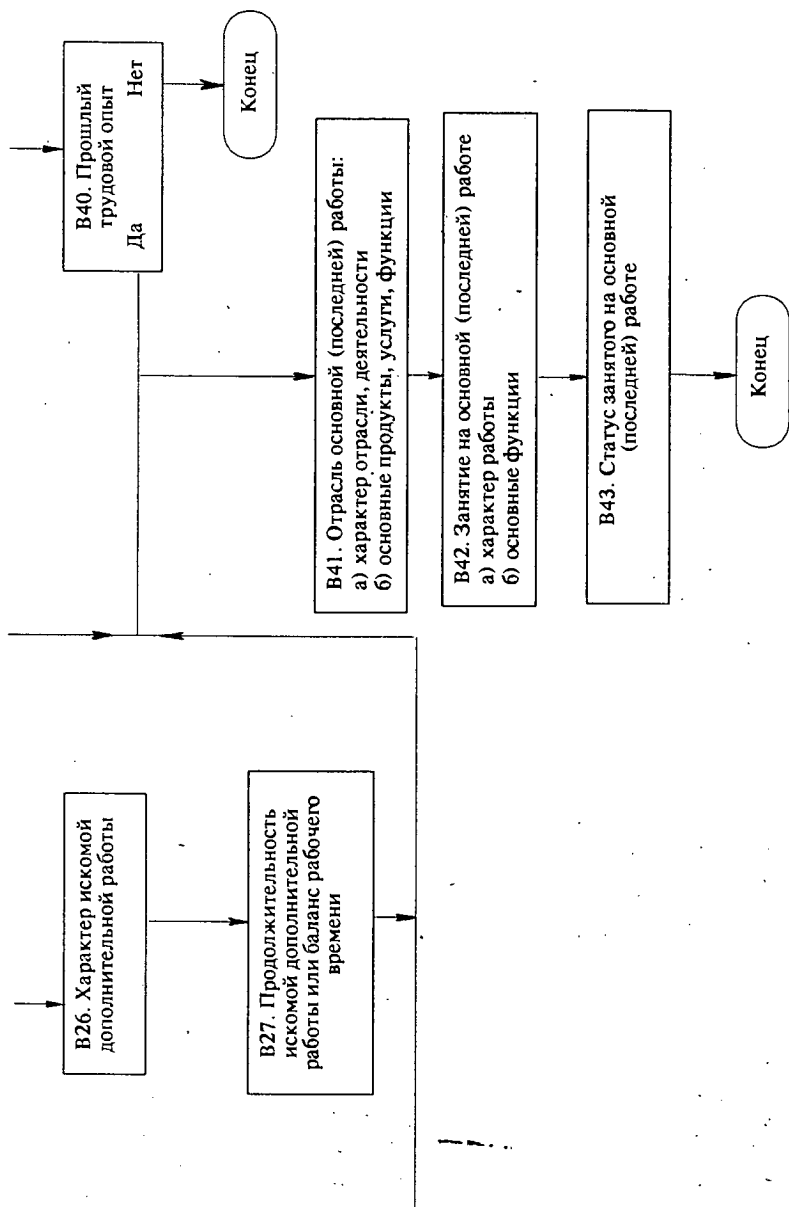
лежащих сбору, с учетом международных стандартов в области статистики экономически активного населения, занятости, безработицы и неполной занятости (ILO, 1983). В ней не перечисляются все возможные вопросы, так как соответствующее число вопросов и их конкретные формулировки определяются применительно к каждой анкете и в зависимости от специфических обстоятельств, которые меняются от страны к стране. Так, в логической схеме 4 о занятости (глава 5) пункт “любая работа” может на деле требовать более одного вопроса для получения необходимой информации (например, отдельных наводящих вопросов для неоплачиваемых членов семьи и работы в течение короткого времени в отчетном периоде). Аналогично, как показано в гл. 2, пункт “перечень видов деятельности” обычно предполагает ряд вопросов о том, выполнялась ли каждая работа, указанная в перечне, во время отчетного периода. В зависимости от обстоятельств вопросы об “отсутствии на работе” и “причинах отсутствия” могут потребовать разных формулировок для различных категорий статуса занятых, например, работающих по найму или на собственном предприятии. Или же может оказаться необходимым предусмотреть специфические вопросы для выявления отдельных групп, например, лиц, которые должны приступить к работе или открыть дело в будущем, увольняемых, сезонных и случайных рабочих и т.п. Цель этих логических схем — проиллюстрировать содержание подлежащей сбору информации по каждой проблеме, без того чтобы приводить все детали, обычно связанные с разработкой реальной анкеты, и уточнять количество и формулировку вопросов на этой стадии разработки.

Связанные в логическую схему 9, схемы призваны также проиллюстрировать еще один полезный инструмент разработки анкеты: графическое представление (сетевая или поточная диаграмма) подлежащей сбору информации, показывающее прохождение опроса, взаимосвязь между вопросами и между подгруппами населения, к которым они относятся. Таким образом, поточная диаграмма становится не только перечнем показателей, а обеспечивает связь между ним и полностью разработанной анкетой. Графическое представление может также быть полезным для уяснения структуры законченной анкеты, включая последовательность вопросов и переходов между ними. Оно помогает также в деле обучения интервьюеров и разработке компьютерных программ для редактирования данных, их перекодирования и табулирования. Еще несколько примеров логических схем анкеты приведены в парагр. 3 настоящей главы. Дополнительные иллюстрации содержатся в работах ООН (United Nations, 1985, p. 174—178) и особенно в (World Fertility Survey, 1976).

Логическая схема 9

Занятость, безработица, видимая неполная занятость и основные экономические характеристики (активность в настоящий момент)





Преобразование показателей в вопросы анкеты

Чтобы построить реальную анкету обследования, понятия, требования к информации и показатели следует превратить в последовательность конкретных, понятных вопросов, на которые можно дать ответы. Главная проблема при этом — степень детализации вопросов.

В этой связи следует указать на общий принцип: единственной гарантией против серьезных пропусков, неполучения ответов и недоразумений является максимальная детализация пунктов и получение информации по каждому пункту или детали, даже если на этапе обработки и анализа некоторые детали будут объединены. Понятия и определения, используемые в обследованиях рабочей силы, часто сложны, и может оказаться необходимым представить какой-либо элемент информации в виде нескольких вопросов не потому, что требуется более детальная информация, а лишь для того, чтобы обеспечить получение ответа, которого ждут от респондента.

Рассмотрим, к примеру, сложное понятие “экономическая активность”, проанализированное в гл. 2. Во многих ситуациях, особенно в развивающихся странах, нет четкой грани между экономической активностью и другими проявлениями активности, например, профподготовкой, добровольным услужением и работой по дому. Кроме того, многие люди могут быть заняты более чем в одном виде деятельности. Традиционный подход, основанный на использовании одного вопроса или небольшой серии вопросов для выявления экономической активности и причастных к ней лиц, может оказаться недостаточным при учете рабочей силы. Более точную и комплексную информацию можно получить при использовании перечня и подробного перечисления всех основных видов деятельности и постановки вопроса о том, причастны ли к ним лица или нет.

Экспериментальные исследования Международного бюро труда, описанные в гл. 2, отражают еще один подход (ILO, 1983 — 1984). В эксперименте, проведенном в Керале (Индия), были использованы две анкеты. Первая призвана была дать информацию о безработице путем постановки прямого вопроса о статусе активности респондентов. Вторая анкета давала ту же информацию на основе ряда более подробных вопросов о видах деятельности респондентов. В результате выявились некоторые любопытные различия между подходами. Применительно к активности в данный момент средний уровень участия в рабочей силе на основе второго, более детального подхода, оказался на пять с лишним процентных пунктов

выше, чем при первой, краткой анкете. Абсолютный разрыв такой же величины (а в относительном выражении он еще больше) отмечен применительно к уровню текущей безработицы, причем расхождения наблюдались во всех половозрастных группах. Однако в отношении *обычной активности* за продолжительный отчетный период оба подхода не дали существенных расхождений как по участию в рабочей силе, так и по показателю безработицы. Вывод на основании этих результатов состоит в том, что более подробный опрос позволяет полнее выявить более зыбкий статус в данный момент, но не дает существенных отклонений применительно к более стабильному отчетному статусу респондентов. Последнее, возможно, отчасти связано с тем, что при выявлении обычного статуса необходимо вспоминать активность за длительный отчетный период.

На практике существуют, однако, пределы возможной детализации вопросов в анкете. Существует много ситуаций, когда практические условия диктуют более упрощенные процедуры и краткие пути, хотя понятно, что более строгие и детальные процедуры, если они возможны, лучше соответствовали бы целям обследования. Желательные детали могут, например, оказаться недоступными из-за невозможности или нежелания респондентов дать соответствующую информацию, общей длины и сложности анкеты, срока проведения работ и ограничений, связанных с аппаратурой ввода и обработки данных.

Главными факторами, определяющими степень детализации содержания обследования в анкете, должны быть цели, предмет и условия проведения обследования. Необходим баланс между требованиями точности и качества данных, с одной стороны, и соображениями стоимости и целесообразности, с другой. Сказанное не означает, конечно, что подробные вопросы всегда гарантируют получение данных высокого качества (хотя часто так и происходит), а краткие формы опросов всегда экономичнее. Чрезмерная детализация на деле способна ухудшить качество полученной информации, не говоря уже о трудностях реализации такого подхода. В то же время видимая краткость анкеты может оказаться весьма обманчивой. Если не добиться нужной детализации и не расшифровать сложные понятия в виде простых вопросов, с виду "краткая" анкета может не только привести к данным низкого качества, но и не дать выигрыша во времени при их сборе. В хороших анкетах детализация вопросов призвана обеспечить не дополнительную информацию, а более точные данные и получение их с меньшими усилиями и даже с меньшими затратами времени.

Эти соображения особенно важны применительно к анкетам обследований рабочей силы. Обследования рабочей силы связаны со сложными понятиями, и их результаты могут существенно зависеть не только от используемых понятий и определений, но и от способа и формы их отражения в анкете. В следующем разделе эти проблемы будут проиллюстрированы на примерах, заимствованных из национальных обследований рабочей силы.

3. НЕКОТОРЫЕ ПРИМЕРЫ ИЗ НАЦИОНАЛЬНЫХ ОБСЛЕДОВАНИЙ РАБОЧЕЙ СИЛЫ

В настоящем параграфе приведены некоторые примеры логических схем анкет из национальных обследований рабочей силы с целью проиллюстрировать различные проблемы, возникающие при определении содержания обследования и его отражении в анкете. В этой связи охарактеризованы некоторые содержательные проблемы, более полно обсужденные в части 1 настоящего руководства. Следует подчеркнуть, что приведенные ниже логические схемы анкеты не представляют собой "образцы", которых необходимо придерживаться. Эти схемы отражают специфические цели обследования, требования к данным и условия проведения работы в отдельных странах, которые могут быть неодинаковыми в разных странах. Надо также отметить, что степень соответствия международным стандартам в примерах различна. Кроме того, приведенные примеры не призваны служить представительной выборкой национальных подходов. Однако они базируются на некоторых сложившихся национальных обследованиях рабочей силы в развивающихся и развитых странах и включают широкий спектр — от краткой анкеты (логическая схема 10) до довольно подробной (логическая схема 13). Эти различия могут отражать действительные или мнимые различия в требующейся информации в соответствующих странах, в ограничениях на затраты и другие моменты. Иногда некоторые виды информации, даже если они желательны, возможно, были исключены из обследования во имя упрощения содержания. Для этой же цели в анкетах иногда выбирают краткий путь, даже если это означает получение менее точной информации, нежели та, которая была бы получена при более полной и детальной последовательности вопросов.

Комментарии будут относиться к четырем группам ключевых показателей: 1) активность и занятость в данный момент; 2) безработица; 3) неполная занятость; 4) обычная активность. В национальных обследованиях рабочей силы можно предусмотреть (и часто предусматривают) также другие показатели, например, характеристики до-

машинных хозяйств (или семей), демографические или социальные характеристики, уровни образования, доход, миграцию и другие характеристики, тесно связанные с занятостью и работой.

Активность и занятость в данный момент

Экономическая активность

Как отмечалось в гл. 2, преобразование понятия экономической активности в соответствующие вопросы — это основная предпосылка точного учета занятости и безработицы в ходе обследований рабочей силы. Однако субъективное толкование этого сложного понятия и отдельных терминов (таких как “экономическая активность” и “работа ради оплаты или прибыли”) интервьюерами и респондентами может отличаться от того, что предполагалось охватить названными понятиями. Проблемы недостаточного отражения экономической активности чаще всего могут возникать в ситуациях, когда значительная часть населения занята деятельностью, отличной от регулярной работы по найму или на своем предприятии, например, случайной или кратковременной работой, работой за плату натурой, помощью в семейном предприятии или производством для собственного потребления. Прежде всего это касается учета экономической активности женщин. В большинстве обследований рабочей силы практика заключается в том, чтобы учесть это с помощью одного или немногих вопросов (обычно основной вопрос строится вокруг ключевого слова “работа”, а один или два уточняющих вопроса относятся к основному виду активности, который в основном вопросе может быть упущен), дополненных в той или иной степени разъяснениями понятия и пограничными примерами в руководстве для интервьюера и в ходе его обучения. Возможно, этого достаточно в обследованиях, проводимых в развитых странах, но в развивающихся странах, где регулярная работа по найму не получила преобладающего распространения, а формы занятости более разнообразны, могут потребоваться более детальные уточнения. Поэтому в гл. 2 приведен альтернативный подход на основе перечня видов активности.

Все логические схемы, представленные в настоящем параграфе, основаны на традиционном подходе, но различаются по уточняющим вопросам, следующим за основным вопросом: относительно работы в течение короткого времени (по меньшей мере в течение одного часа в отчетном периоде) — в схеме 10; отсутствие уточнений — в схеме 11, уточнение относительно работы на собственной (семейной) ферме — в схеме 12; уточнение относительно неоплачиваемой работы на семейном предприятии — в схеме 13. Каждое из

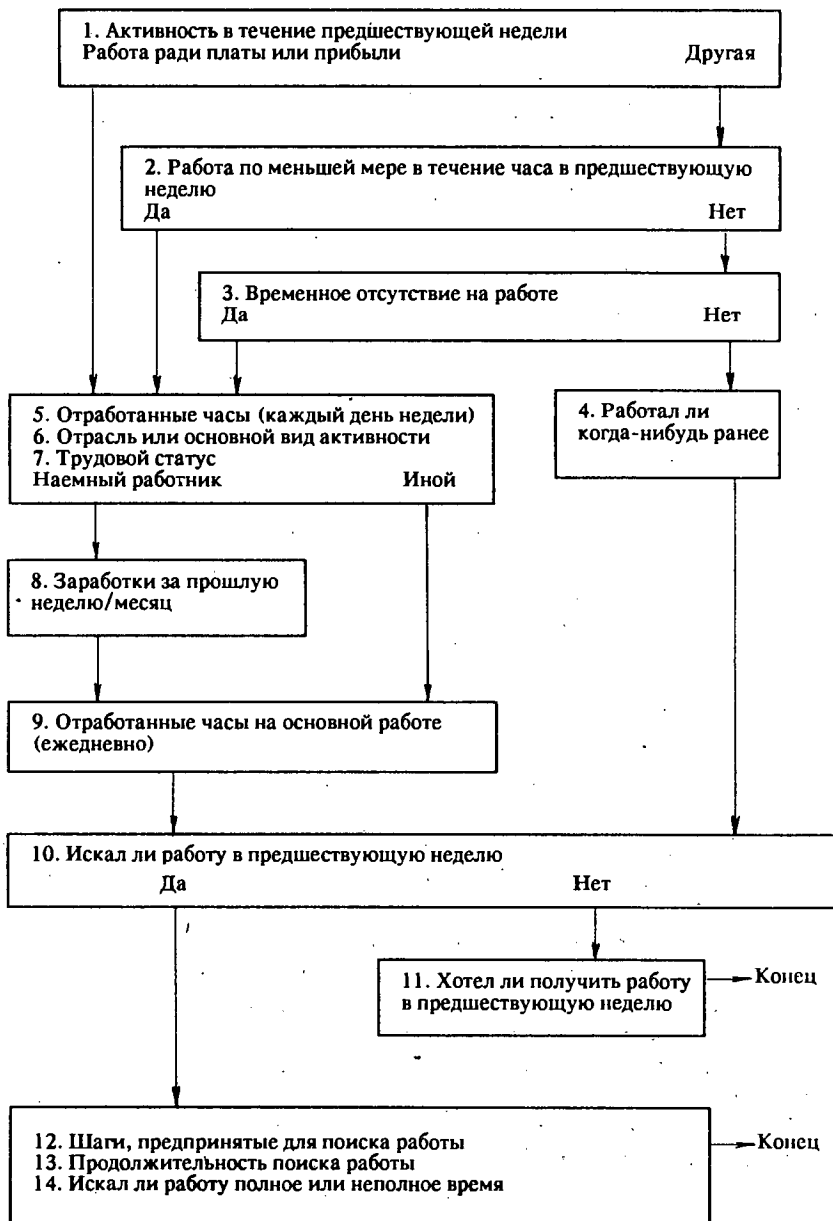
этих уточнений может быть полезным для выявления определенных форм экономической активности, отличных от регулярной работы по найму. В зависимости от ситуации в области занятости в данной стране может, однако, быть целесообразным включение в анкету нескольких уточняющих вопросов, особенно когда невозможно включить полный перечень видов активности.

При сопоставлении логической схемы 10 с логической схемой 12 видно, что важное различие между ними состоит в том, что лишь вторая содержит в явном виде вопрос о работе на собственной (семейной ферме). И это несмотря на то, что обе схемы относятся к развивающимся странам с большой долей населения, занятого на собственных сельскохозяйственных предприятиях. Учитывая условия страны, представляется, что схема 10 скорее "ориентирована на город" — обстоятельство, которое, похоже, часто встречается в национальных обследованиях рабочей силы. Заметим, что вопрос о работе на собственной (семейной) ферме в схеме 12 рассчитан на две подгруппы для различных целей: а) применительно к занятым лицам он имеет целью определить последовательность опроса, т.е. пропустить несколько вопросов (о занятости, подготовке, доходе и т.д.), так как они считаются несущественными (или неважными) для лиц, занятых на собственной (семейной) ферме; б) применительно к лицам, которые не отнесены к занятым в результате предшествующих вопросов, этот вопрос преследует уточняющую цель и позволяет полнее учесть неоплачиваемых членов семьи и лиц, занятых производством для собственного потребления.

Временное отсутствие на работе

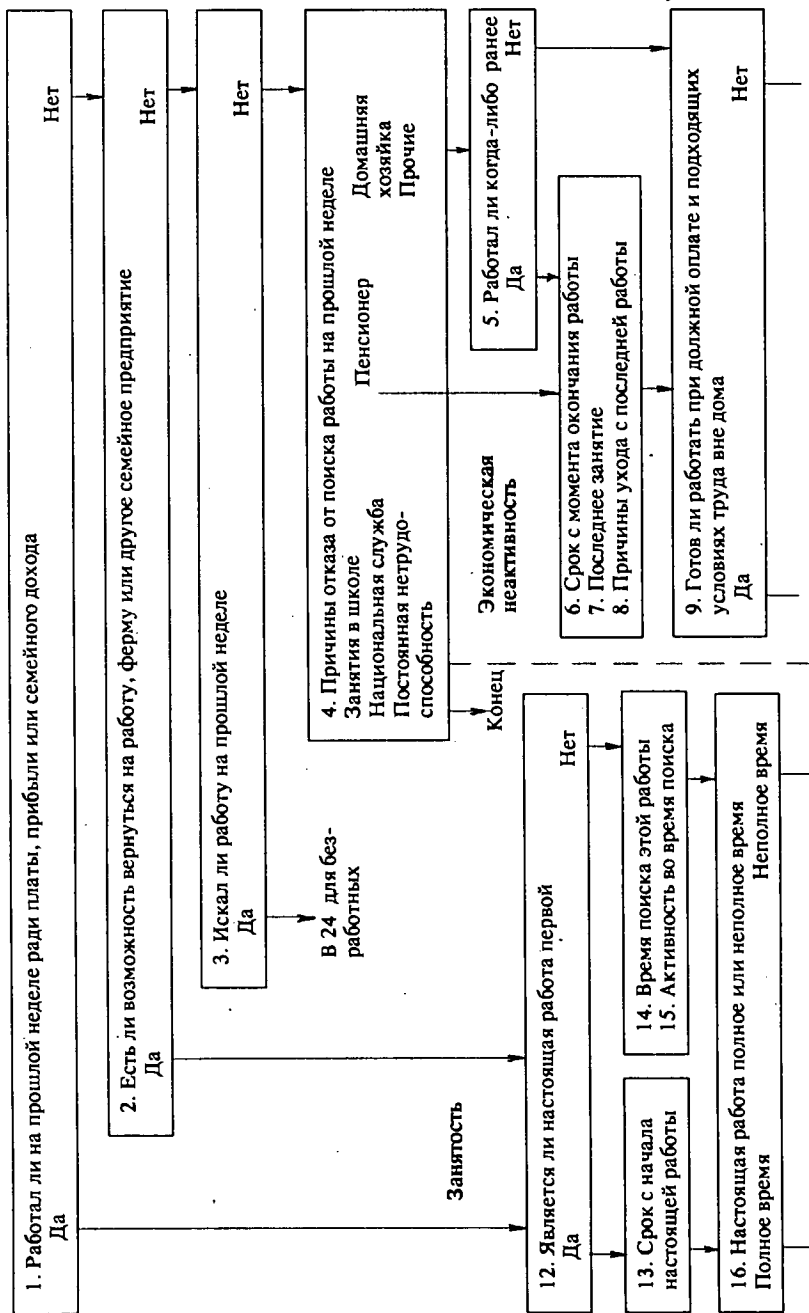
Как отмечалось в гл. 5, занятые включают не только работающих лиц, но и тех, у кого есть работа или предприятие, но они временно отсутствовали на работе в отчетном периоде. В международных стандартах установлен ряд критериев для уточнения временного отсутствия на работе наемных лиц и занятых на собственном предприятии. Применительно к наемным работникам эти критерии основаны на понятии формальных связей с предприятием в виде получения заработной платы, заверения о возвращении на рабочее место (или соглашения о дате возвращения) или окончания срока отсутствия на работе. Применительно к занятым на собственных предприятиях эти критерии проще. На практике для работодателей и владельцев мелких фирм без наемных лиц временное отсутствие на работе подтверждается дальнейшим существованием предприятия и учетом продолжительности отсутствия на работе.

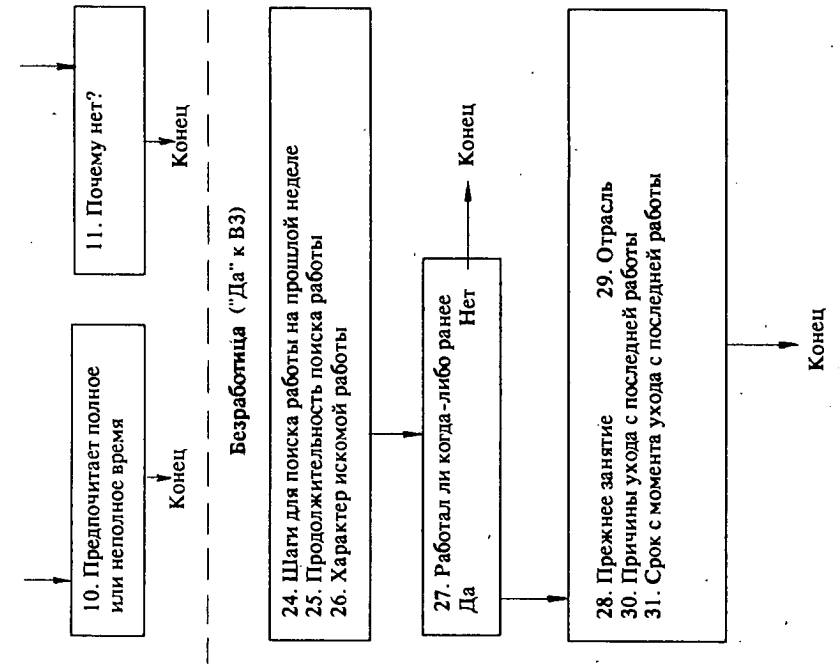
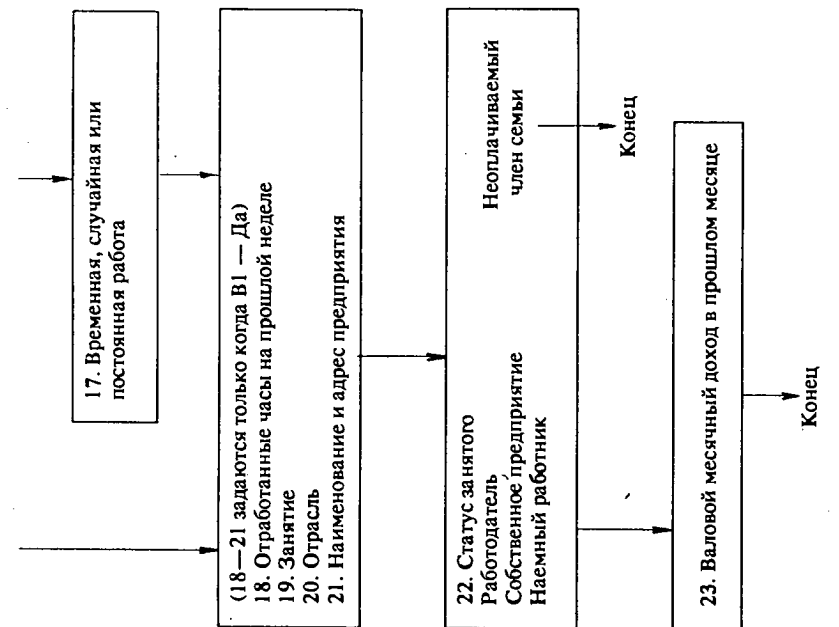
Логическая схема 10
Анкета национального обследования рабочей силы: пример 1



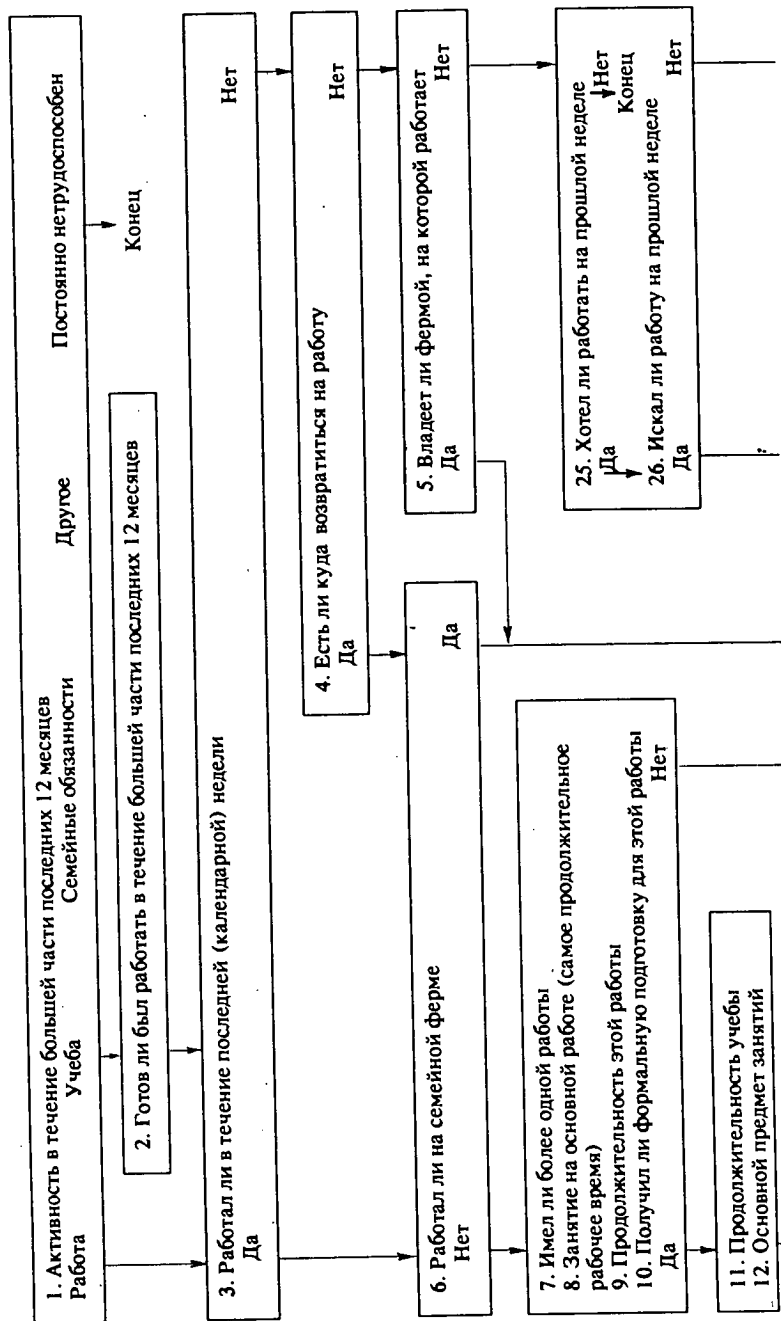
Логическая схема 11. Анкета национального обследования рабочей силы: пример 2

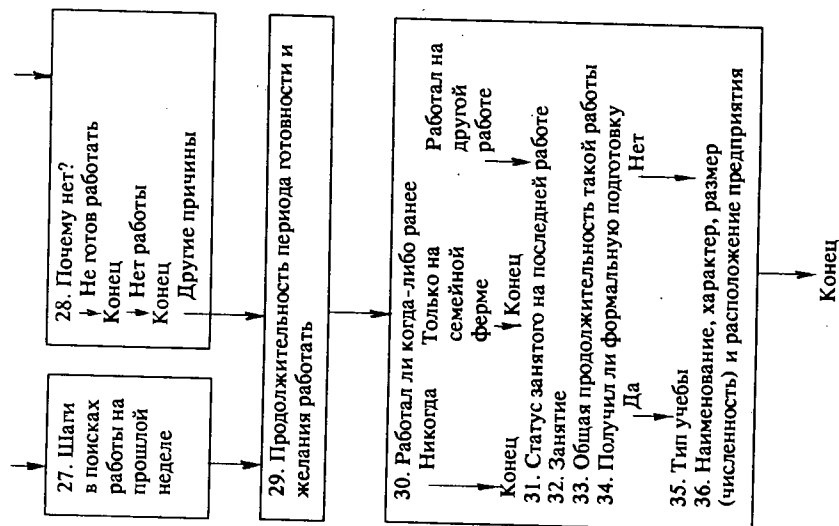
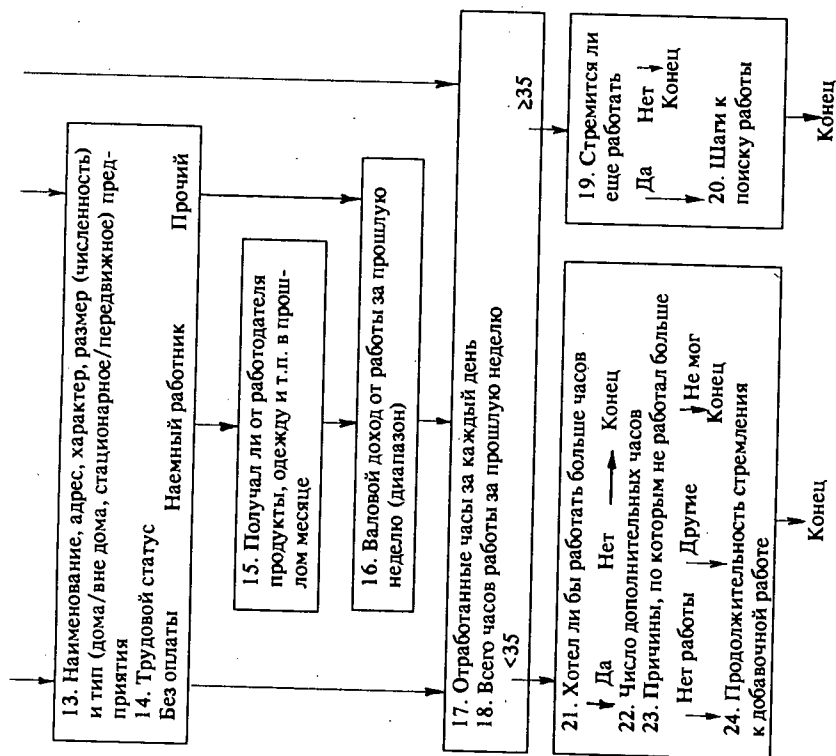
Экономическая активность



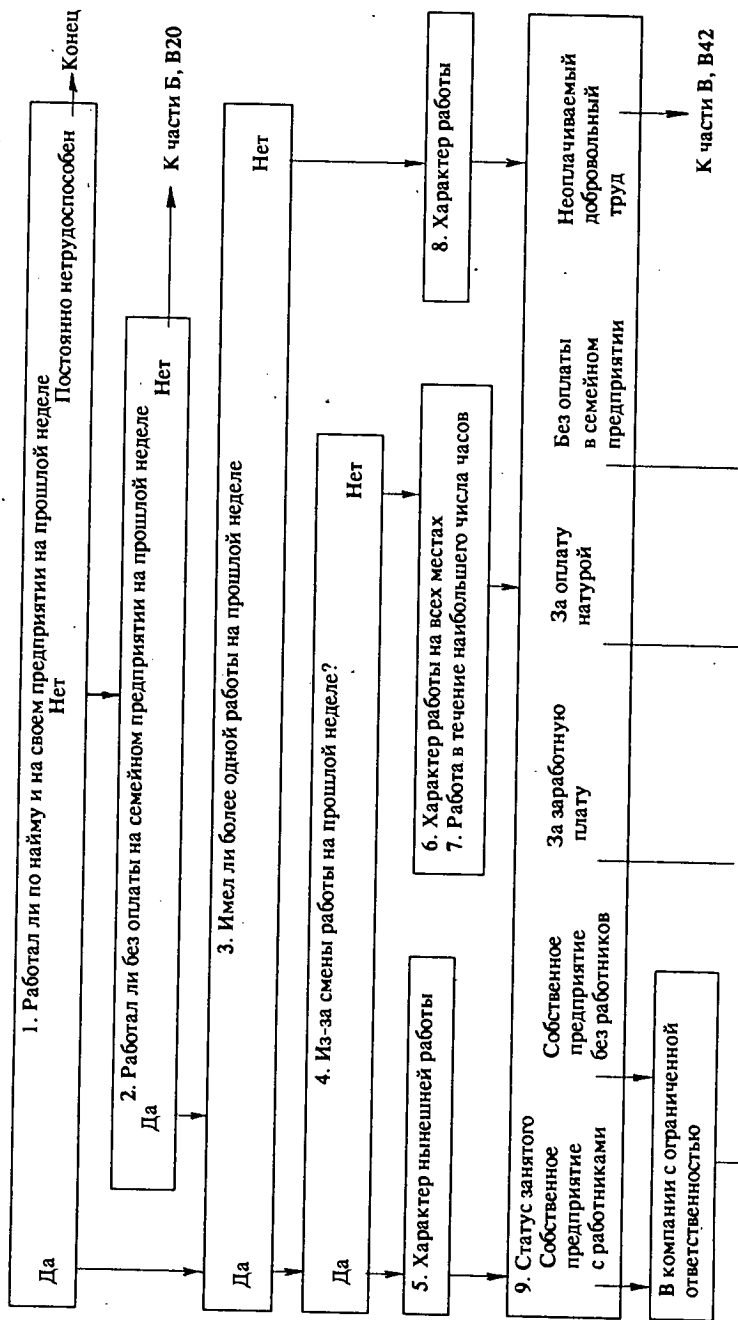


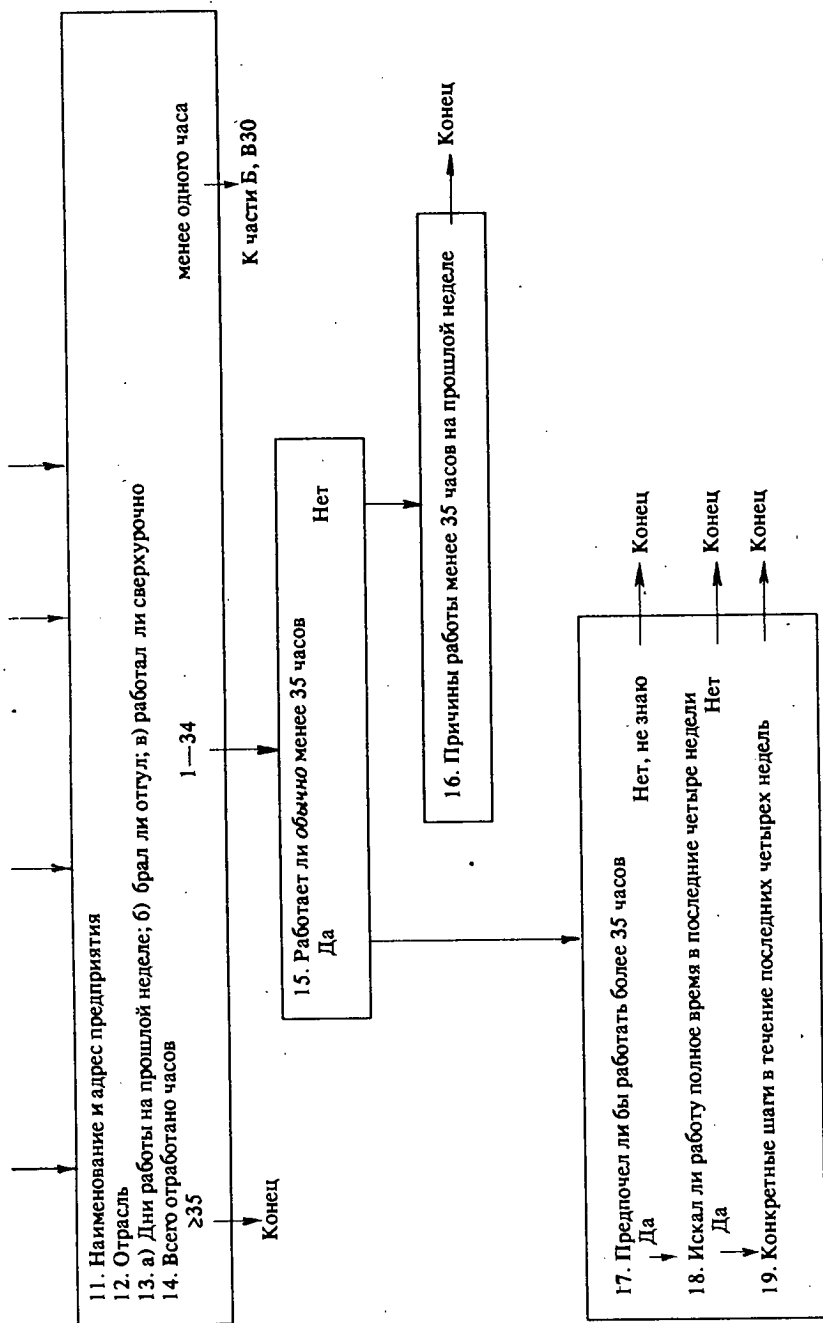
Логическая схема 12. Анкета национального обследования рабочей силы: пример 3





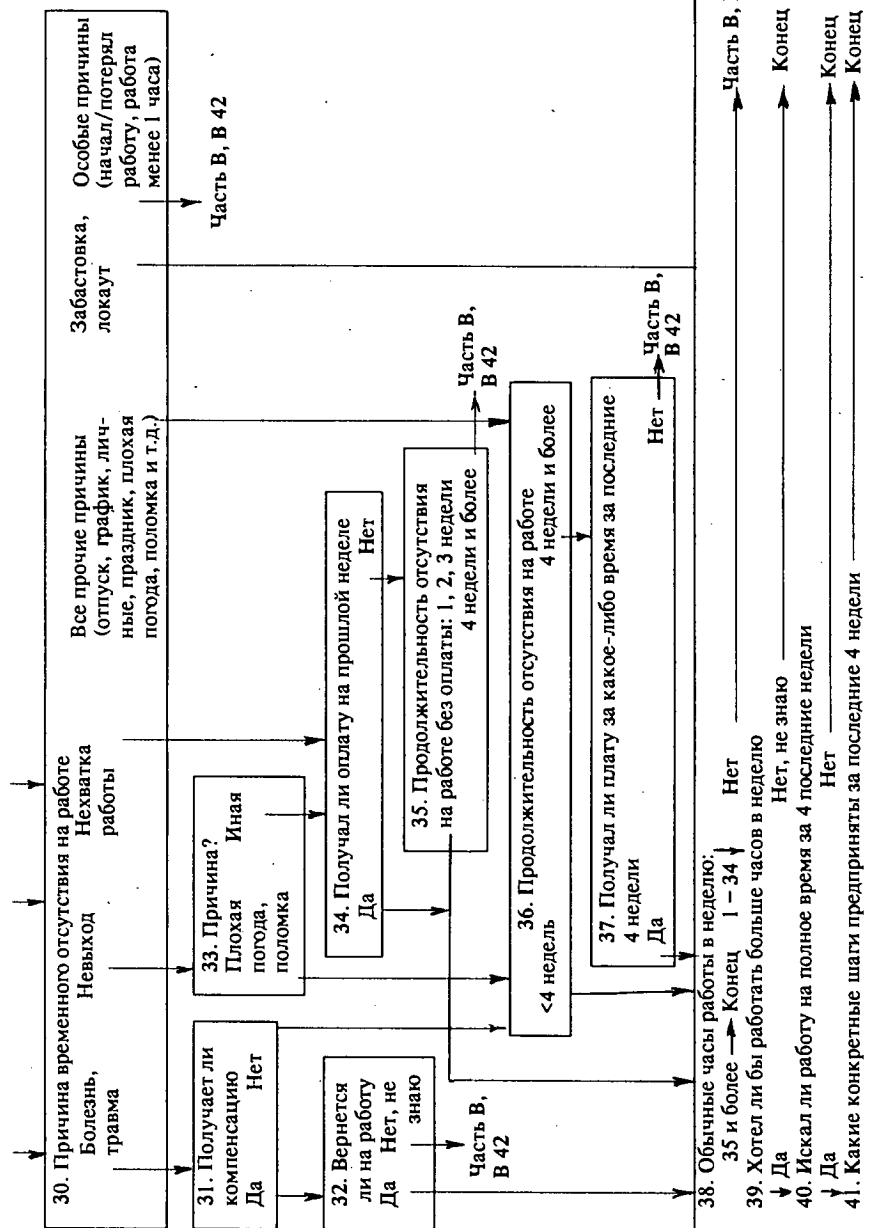
Логическая схема 13. Анкета национального обследования рабочей силы: пример 4
ЧАСТЬ А. Лица, фактически работавшие в неделю обследования



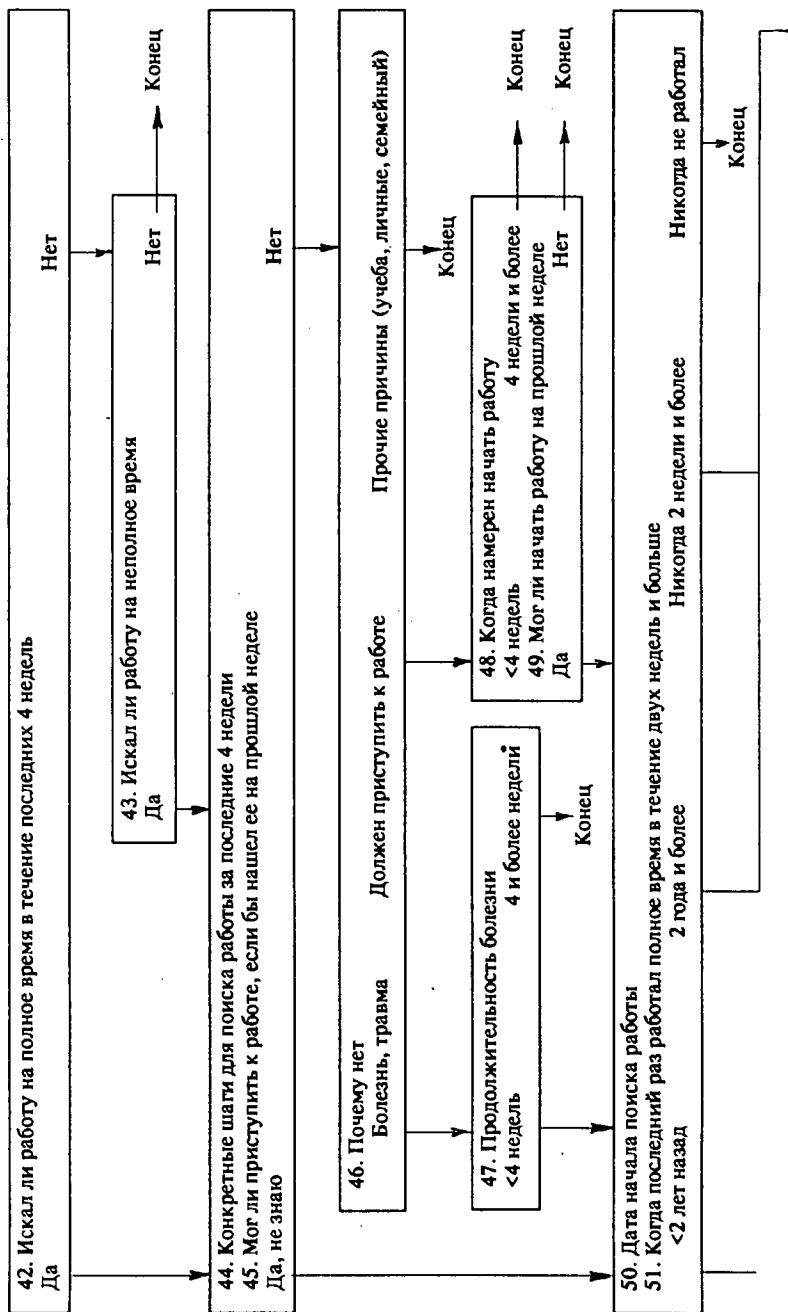


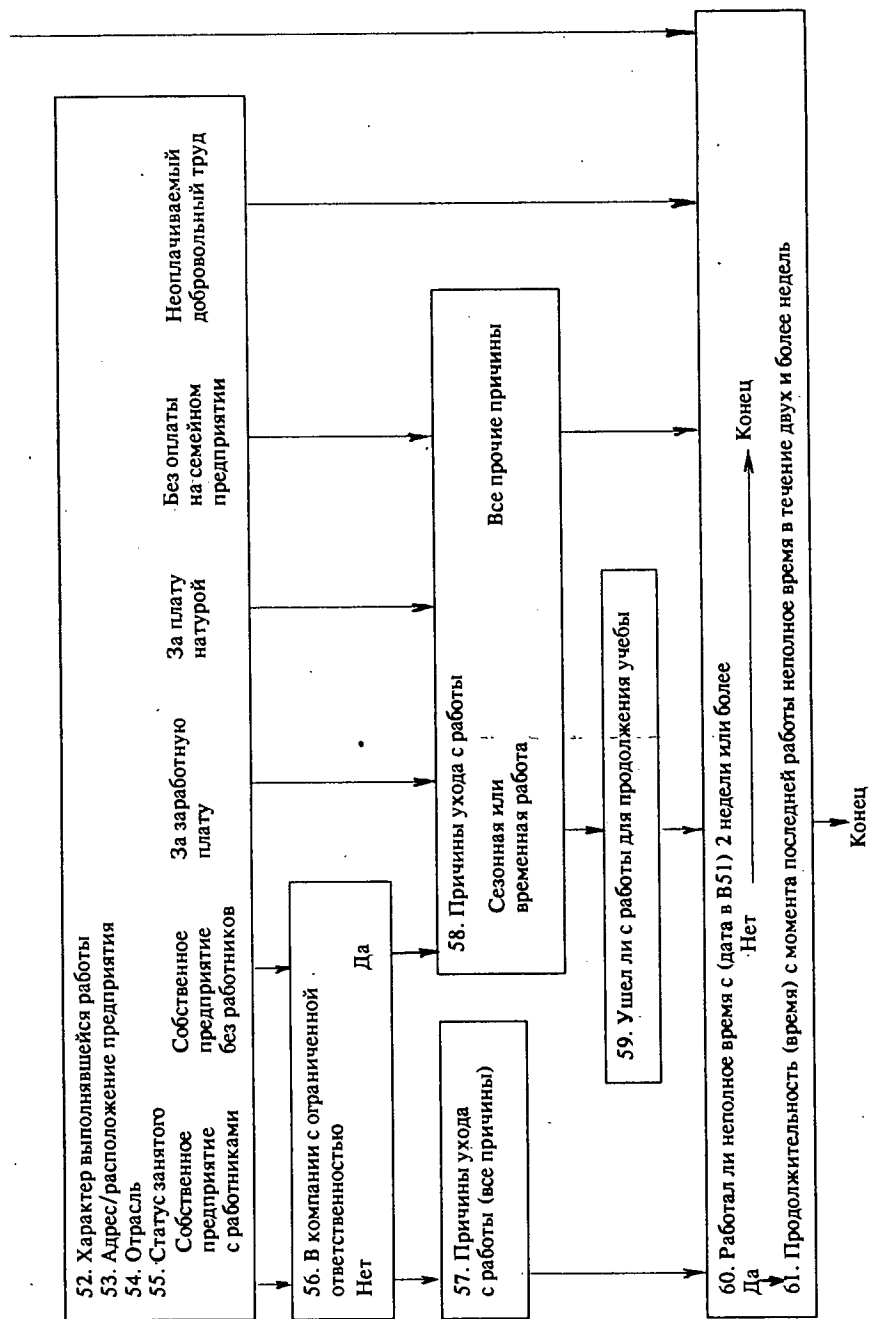
ЧАСТЬ Б. Лица, имевшие место работы, но не работавшие

20. Имел ли работу, на которой отсутствовал на прошлой неделе из-за отпуска, болезни или другой причины Да Нет		К части В, В 42	
21. Имел ли более одной работы на прошлой неделе Да Нет		Нет	
22. Характер работы на каждом месте 23. Работа с наибольшим числом часов		24. Характер выполнявшейся работы	
25. Место работы (адрес); 26. Отрасль; 27. Статус занятого			
За заработную плату	За плату натурой	Собственное предприятие с работниками	Собственное предприятие без работников
28. В компании с ограниченной ответственностью Да Нет		Без оплаты на семейном предприятии	
Нет		Неоплачиваемый добровольный труд	
Да		К части В, В 42	
29. Причины отсутствия на работе (все причины)			



ЧАСТЬ В. Поиск работы и участие в рабочей силе





То, какое внимание следует уделить в данной анкете точному перечислению категории “временно отсутствующих на работе”, зависит от величины и значимости этой группы, а также расходов и сложности анкеты, связанных с таким перечислением. В логической схеме 4 (гл. 5) в явном виде предусмотрено включение, по сути дела, двух видов информации: имел ли индивид работу или собственное предприятие, где он временно отсутствовал, и причина отсутствия на основе специального перечня. В явном виде не предусмотрены какие-либо критерии “сохранения формальной связи с предприятием” или “продолжения существования предприятия”, так как исходили из того, что во многих странах будет достаточно предложить в руководстве для интервьюеров уточнение этого вопроса. Помимо прочего, в руководстве необходимо пояснить, что в эту категорию не входят лица, которые должны приступить к работе лишь в будущем, неоплачиваемые члены семьи, случайные рабочие и вообще все неработающие, которые не сохраняют формальной связи с предприятием или чье предприятие прекратило существование во время их отсутствия. Но, как отмечалось в гл. 5, дополнительные вопросы могут понадобиться в тех случаях, когда преобладающая структура занятости делает необходимым более точный учет этой категории лиц.

Последовательность вопросов в логических схемах 10—12 короче и менее точна, чем в логической схеме 4. Включен лишь один вопрос типа “имеется ли у индивида рабочее место или предприятие, куда он может вернуться”, однако причины временного отсутствия не выясняются. В гл. 5 показано, что добавление вопроса о причинах отсутствия может быть полезным в нескольких отношениях, включая возможность уточнения предыдущего вопроса, а тем самым лучшего уяснения обстоятельств, которые оправдывают выделение этого момента в самостоятельный признак. Если равноценную информацию не получить на основе других разделов анкеты, то отсутствие такого вопроса приводит к приблизительным результатам, значение этой неточности зависит от численности данной категории лиц в стране.

Напротив, в логической схеме 13 содержится значительно более подробная серия вопросов по этой проблеме по сравнению с логической схемой 4. Во-первых, формулировка основного вопроса (имеется ли рабочее место или предприятие, на котором индивид отсутствовал в данный момент) уже предполагает возможные причины (отпуск, болезнь и т.д.), а это помогает респонденту уяснить себе понятие. Во-вторых, причины временного отсутствия увязаны со статусом (что вытекает из международных стандартов) благодаря

отдельным перечням причин для: а) наемных работников и владельцев компаний с ограниченной ответственностью; б) прочих владельцев и неоплачиваемых членов семьи; те, кто занят на неоплачиваемых добровольных работах, в явном виде исключены. В-третьих, подгруппе (а) задают особые вопросы в зависимости от причин отсутствия на работе, с учетом критериев наличия формальной связи с предприятием, а именно, получения компенсации или заработной платы во время отсутствия, продолжительности отсутствия, надежды вернуться на то же предприятие. Очевидно, что такой подход предполагает довольно сложную последовательность вопросов с переходами, поэтому необходимо сопоставить затраты с пользой, которую можно получить от этой информации.

Безработица

Согласно определению международных стандартов (см. гл. 6), безработные — это лица старше определенного возраста, которые: а) не имели работы в течение отчетного периода; б) готовы сейчас приступить к работе и в) ищут работу, т.е. в течение определенного прошлого периода (например, в последние четыре недели) предпринимали конкретные шаги для получения работы по найму или создания собственного предприятия. Критерий готовности относится к способности и желанию работать, если такая возможность появится. Говоря о критерии поиска работы, важно охватить не только тех, кто ищет работу по найму, но и тех, кто стремится основать собственное дело. Стандартное определение безработицы распространяется и на лиц, не имеющих работы и готовых сейчас приступить к ней, которые договорились о начале работы или открытии своего предприятия после отчетного периода, независимо от того, продолжают ли они поиск работы. Кроме того, лица, временно уволенные работодателем (без сохранения формальной связи с предприятием) и готовые сейчас приступить к работе, могут быть отнесены к безработным, даже если они активно не ищут работу. Очевидно, что степень детализации в анкете рабочей силы (и/или в руководстве для интервьюеров) всех сложных критериев и различных ситуаций безработицы зависит от конкретных обстоятельств и требований к данным.

Еще один момент, который особенно относится к развивающимся странам, связан с возможным использованием видоизмененного определения безработицы, при котором ослаблен критерий поиска работы. Это предусмотрено в международных стандартах для ситуаций, когда традиционные способы поиска работы имеют мало

смысла, когда рынок труда в целом неорганизован или не получил развития, способность рынка поглотить рабочую силу ограничена или же распространена работа на собственных предприятиях. На практике степень смягчения критерия поиска работы может быть разной и относиться ко всем или некоторым категориям трудящихся (например, сезонным рабочим, ожидающим начала сезона, "обескураженным трудящимся" и т.п.). Как отмечено в гл. 6, при ослаблении критерия поиска критерий готовности приобретает решающее значение для учета безработицы, и его следует уточнять как можно полнее с точки зрения желания работать, прежнего опыта работы, приемлемой работы и т.п.

Логическая схема 6 в гл. 6 призвана вобрать эти различные требования. Она содержит вопросы о желании работать, шагах по поиску работы, причинах отказа от поиска и готовности сейчас приступить к работе, дополненные уточняющими вопросами о готовности к работе, причинах неготовности и прошлом опыте работы. Вопрос о причинах отказа от поиска работы помогает выявить не только тех, кто приступает к работе в будущем, кто временно уволен, но и другие группы лиц, отказавшиеся от поиска работы. Следовательно, логическая схема позволяет ввести различную степень ослабления критерия — от полной его отмены до последовательного применения.

Соответствующие серии вопросов в логических схемах 10—13 в той или иной степени отличаются от логической схемы 6 в гл. 6. Например, в логических схемах 10 и 11 информацию о готовности сейчас приступить к работе получают лишь от тех, кто не ищет ее, в предположении, что лица, которые активно искали работу в течение отчетного периода (прошлой недели), готовы приступить к ней. Но это может оказаться неверным в отношении лиц, ищущих работу, которую следует начать лишь на более позднюю дату. В логической схеме 10 вопрос сформулирован несколько неопределенно в смысле желания в настоящее время работать, тогда как в логической схеме 11 характер приемлемой работы уточнен в формулировке вопроса и добавлен еще один вопрос о предпочтительности работы полное/ неполное время в качестве еще одного теста готовности к работе.

В логических схемах 10 и 11 причины отказа от поиска либо вовсе не перечислены, либо даны в форме, не позволяющей выявить лиц, находящихся в особых обстоятельствах (приступят скоро к работе, временно уволены), или ослабить критерий поиска, что может быть целесообразным. Кроме того, ни основной вопрос, ни заранее перечисленные варианты ответов (в примере они не показаны) на вопрос о "шагах по поиску работы" не охватывают ситуацию, когда не

ищут работу по найму, а пытаются создать собственное предприятие. В результате этих изъянов может возникнуть серьезный недостаток отдельных групп, например, желающих основать собственное дело и отказавшихся от поиска работы по уважительной причине. Обе эти группы могут быть многочисленными и иметь важное значение в развивающихся странах.

В логической схеме 12 серия вопросов больше похожа на логическую схему 6, но она несколько проще. Основное различие состоит в том, что отсутствуют явные вопросы о готовности работать *сейчас*, хотя есть вопрос о “продолжительности периода, когда индивид был готов и хотел работать”; скорее информация о готовности (фактически, о неготовности) работать выступает как единственный вариант ответа на вопрос о причинах отказа от поиска работы. Это еще один пример “срезания углов” в анкете, которое, видимо, было признано приемлемым с учетом национальных условий и потребности в данных.

Общей чертой логических схем 10-12 является то, что для таких аспектов, как активный поиск, готовность и желание работать взят тот же отчетный период (одна прошлая неделя). Такой выбор, возможно, упростил опрос, но, как поясняется в части 1 настоящего руководства, для этих различных видов показателей часто более целесообразны разные отчетные периоды. Такой подход детально пытаются реализовать в логической схеме 13. Этими периодами являются: последние четыре недели — для поиска работы; последняя неделя — для готовности к работе; ближайшие четыре недели — для начала работы. Те, кто был слишком болен или травмирован, чтобы приступить к работе в течение четырех недель, также выделены отдельно. Это привело к довольно разработанной серии вопросов. Еще одна особенность логической схемы 13 заключается в том, что в отличие от логической схемы 6 в ней отсутствует вопрос о причинах отказа от поиска работы и проверка готовности приступить к ней. Можно предположить, что эти вопросы не очень нужны, когда критерий поиска строго применяется и не делается исключений для “обескураженных трудящихся” и аналогичных категорий.

Неполная занятость

В гл. 7 отмечается, что следует различать две основные формы неполной занятости: видимую неполную занятость, отражающую недостаточный объем занятости, и невидимую неполную занятость, которая характеризуется такими моментами, как низкий доход, низкая производительность и неудовлетворительное использование

квалификации трудящихся. При разработке анкеты обследования рабочей силы необходимо принять решение о том, сколько информации о неполной занятости и в какой форме необходимо (и на деле можно) получить с учетом условий данной страны. Учет невидимой неполной занятости при обследованиях рабочей силы может быть сопряжен с особыми трудностями. Признавая эти трудности, в международных стандартах отмечается, что по техническим причинам статистический учет неполной занятости можно ограничить ее видимой частью. Однако и учет видимой неполной занятости, хотя и более простой, также требует постановки ряда вопросов, о чем уже говорилось в гл. 7.

В международных стандартах предусмотрены два элемента учета видимой неполной занятости: число неполностью занятых лиц и объем неполной занятости. Неполностью занятых лиц выявляют на основе трех критериев: а) работа меньше нормального рабочего времени; б) вынужденный характер такой работы; в) поиск работы и готовность приступить к ней в течение отчетного периода. Названные критерии относятся к занятому населению в целом, т.е. к лицам наемного труда и занятым на собственных предприятиях, работающим в данный момент и временно отсутствующим на работе. Чтобы учесть аспект (а), проще всего было бы поставить перед респондентами прямой вопрос, работают ли они полное или неполное время и работали ли они меньше нормального времени. Но чтобы получить более точные данные, может потребоваться обходной путь, особенно в развивающихся странах, где рабочее время большинства населения не регламентируется в договорном порядке. Время, отработанное в отчетном периоде, и нормальное рабочее время следует получить отдельно и сопоставить. Что подразумевать под "нормальным временем" — это сложный вопрос, и выявлять его в каждом отдельном случае при обследовании может оказаться затруднительным. Поэтому на практике может оказаться необходимым прибегнуть к весьма упрощенному подходу, например, единой норме для всех видов деятельности и всех категорий трудящихся. Придерживаясь этого подхода, необходимо позаботиться о том, чтобы к числу неполностью занятых не были отнесены лица, которые заявляют о работе меньше нормального времени, но фактически являются полностью занятыми, так как их полное рабочее время не предполагает большого количества часов (например, учителя и судьи). Кроме того, при определении общего количества отработанных часов за отчетную неделю следует учитывать совместительство, а также необходимость отдельно считать часы работы за каждый день, если отчетным периодом служит целая неделя. Чем разнообразнее и

изменчивее ситуация, тем тщательнее надо подходить к постановке вопросов, чтобы получить достоверную информацию. Применение критерия (б) — вынужденного характера работы неполное время — обычно требует оценки причин работы меньшего количества часов (или вообще невыхода, если речь идет о временно отсутствующих на работе). Эти причины можно выявить на основе спонтанных ответов респондентов или же основные причины можно указать в перечне. В отдельных случаях может оказаться необходимым дополнительное уточнение с целью выяснить, является ли причина добровольной или вынужденной. Критерий (в) — поиск или готовность к дополнительной работе — в рамках измерения неполной занятости рассчитан прежде всего на лиц, вынужденных работать неполное время. Следовательно, вопросы в этом случае могут быть проще, чем при изучении этого аспекта при учете безработицы. Масштабы (объем) видимой неполной занятости — это суммарное время возможной дополнительной работы в отчетном периоде применительно к каждому неполностью занятому лицу. Поэтому для оценки этого объема требуется, помимо выявления неполностью занятых лиц, получить информацию о дополнительной работе, которую ищет и к которой готов приступить индивид, чтобы работать нормальное время.

Из сказанного выше очевидно, что точный учет неполной занятости предполагает достаточно сложный комплекс вопросов, поэтому на практике упрощения могут быть неизбежными. Приведенные в настоящем разделе четыре логические схемы более или менее отличны друг от друга и от схемы, представленной в гл. 7. Совершенно отсутствуют вопросы о видимой неполной занятости в логической схеме 1, относящейся к стране, где эта проблема, очевидно, не имеет большого значения. В логической схеме 10, относящейся к развивающейся стране, где проблема неполной занятости имеет более важное значение, вопросы включают фактически отработанные часы (отдельно за каждый день отчетной недели, чтобы добиться большей точности), информацию о поиске или желании дополнительной или другой работы — полное или неполное время, а также конкретные шаги по поиску такой работы. Отсутствуют вопросы о причинах работы неполное время или невыхода на работу в отчетном периоде, т.е. предполагается, что все, кто ищет или стремится к дополнительной или другой работе, были заняты неполное время вынужденным образом. Информация о продолжительности искомой дополнительной работы сводится к довольно общему вопросу о том, ищет ли индивид работу на полное или неполное время. Вопросы относятся к тем, кто ищет работу по найму, и достаточно учитывают занятых

на собственных предприятиях. Логическая схема 12, которая также относится к стране, где неполная занятость — это серьезная проблема, включает ряд вопросов, больше похожих на вопросы схемы в гл. 7. Исходя из нормального рабочего времени в 35 часов в неделю, схема конкретно адресуется тем, кто работал меньше в течение отчетной недели, с вопросами: хотели ли бы они работать больше, о желательном добавочном количестве часов и причинах, по которым они не работают больше часов. Таким путем выясняют, является ли такой режим вынужденным, и получают объем видимой неполной занятости. Те же вопросы адресованы тем, кто временно отсутствует на работе. Тех, кто работал полное время (35 часов и более) в течение отчетной недели, спрашивают, ищут ли они другую работу и какие шаги они для этого предпринимали, очевидно, с тем чтобы получить приблизительную оценку невидимой неполной занятости. Наконец, в логической схеме 13 интересуются причинами работы меньше нормального времени (в качестве нормы взяты 35 часов) в отчетном периоде только у тех, кто обычно занят 35 часов и более, в предположении, что те, кто фактически и обычно работал менее 35 часов и искал дополнительную работу, делали это вынужденно. В то же время критерий поиска дополнительной работы не проверен применительно к тем, кто фактически отработал менее 35 часов, но обычно работает больше; сказанное относится и к лицам, обычно работающим 35 часов и более, но временно отсутствующим на работе. В отличие от логической схемы 12, вопросы не позволяют (и, похоже, так было задумано) получить точную оценку объема видимой неполной занятости.

Обычная активность

Приведенные выше рассуждения относятся преимущественно к активности в определенный момент в течение краткого отчетного периода. Но, как отмечалось в гл. 4, в странах, где наблюдаются значительные сезонные и другие колебания экономической активности, важно получить информацию о преобладающей структуре активности в течение более продолжительного периода, например, года. Учет обычной активности в течение продолжительного периода, как правило, сопряжен с более серьезными проблемами, чем моментальная фотография активности в течение короткого отчетного периода. Существует выбор между двумя основными подходами: а) повторять текущие опросы во времени, с тем чтобы охватить период нужной продолжительности; б) использовать сам продолжительный период в качестве отчетного с ретроспективной оценкой.

Вариант (а) часто неприемлем из-за большого объема работ и значительных расходов. Потребность в неодинаковых опросах и последующем объединении информации по каждому индивиду за несколько раундов обследования также ведет к усложнению плана и проведения работы. В то же время вариант (б) может быть связан с серьезными ошибками в ответах из-за длительного периода, к которому относятся воспоминания, а также из-за многообразия видов деятельности индивида в течение года. Такова одна из причин, по которой использование этих подходов может дать разные результаты. Последовательность вопросов, необходимых для точной ретроспективной оценки обычной активности, может оказаться довольно сложной, а это в свою очередь потребует весьма тщательного построения и оценки анкеты. На практике во многих обследованиях рабочей силы и переписях населения вынуждены прибегать к более простым и кратким сериям вопросов для учета обычной активности, что объясняется практическими трудностями, присущими продолжительному отчетному периоду.

При ретроспективных опросах точность и полезность информации можно повысить, разбив весь отчетный период на соответствующие отрезки или подпериоды и задавая вопросы для каждого из них. Для такого подхода имеются веские основания, хотя он и удлиняет опрос. Во-первых, постановка вопросов отдельно для каждого отрезка времени помогает вспомнить события. Например, как отмечено в гл. 4, полумесячные воспоминания (в отличие от воспоминаний для года в целом) помогают уменьшить ошибки. Аналогичным образом, применительно к сельскохозяйственной деятельности качество данных можно улучшить, если задавать вопросы отдельно по каждому сезону сельскохозяйственных работ и межсезонью. Во-вторых, разбивка отчетного периода на более короткие отрезки для опроса позволяет точнее выявить основной статус активности.

Следует напомнить, что в рамках учета обычной активности индивидов прежде всего подразделяют на обычно неактивных и обычно активных в зависимости от того, был ли индивид "занят или безработным" в течение большей части (доля эта заранее определяется) продолжительного отчетного периода; только потом, и если это необходимо, обычно активных делят на занятых и безработных с учетом преобладающего вида активности. Очевидно, что результаты могут существенно различаться в зависимости от выбора единицы времени для определения обычного статуса (например, день, неделя, полмесяца или месяц) и от того, насколько разнообразен характер активности данных лиц. В большинстве случаев маловероятно, чтобы одного общего вопроса об основной активности было

достаточно для правильного и единообразного применения понятия активности.

Понятие "обычной активности" включает в себя как статус занятого (включая временное отсутствие на работе), так и безработицу. Сочетание этих двух состояний и определяет, следует ли индивида относить к числу обычно активных в течение продолжительного отчетного периода. Это означает, что задавая отдельно (как в логической схеме 12) вопросы о том, был ли индивид: а) занят большую часть отчетного периода или б) безработным большую часть этого периода, не позволяет полностью выявить обычно активное население, так как такой статус определяется сочетанием этих двух состояний.

Решение указанных проблем заключается в том, чтобы выявить продолжительность и периоды занятости, безработицы и экономической неактивности в такой форме, при которой они не пересекались бы и в сумме давали общую продолжительность отчетного периода. Пример такого подхода на основе членения года на периоды для облегчения воспоминаний приведен в гл. 4.

4. ПРИНЦИПЫ И МЕТОДЫ ПОСТРОЕНИЯ АНКЕТЫ

Структурирование опроса

Важнейшая проблема при построении анкеты заключается в выборе меры структурирования опроса. Максимальная структуризация предполагает использование заранее определенных и единообразных процедур проведения опроса без всяких отклонений в выборе последовательности и формулировки вопросов, форме и вариантах ответов и их кодирования. Интервьюер обязан строго придерживаться формулировок вопросов анкеты, инструкций о том, как и кого опрашивать, а также того, какие возможны варианты ответов на каждый вопрос.

Высокая степень структуризации опроса может быть желательной во многих случаях, например, когда обследование включает сложные вопросы, ответы на которые зависят от точки зрения респондента и от того, насколько точно сформулирован вопрос. Структурирование может оказаться особенно желательным при крупномасштабных обследованиях по нескольким причинам. Во-первых, такие обследования требуют привлечения значительного числа работников, чьи способности и квалификация не могут быть одинаковыми и одинаково высокими, поэтому необходимо предусмотреть стандартные процедуры. Во-вторых, структурированный опрос позволяет получить данные, которые легче кодировать и табулиро-

вать, а это соображение может оказаться крайне существенным при крупномасштабных обследованиях, в ходе которых получают огромные массивы данных. В-третьих, структурирование и контроль позволяют получить сопоставимые результаты в ходе непрерывных обследований. Все это обычно важно при общенациональных и крупномасштабных обследованиях рабочей силы.

Однако максимальная структуризация и стандартизация — это не всегда самый удачный подход. Имеются случаи, когда желательно предоставить интервьюеру свободу выбора способа подачи вопросов и формы регистрации ответов. Например, условия, язык и другие моменты опроса могут быть столь различными, что заранее определить стандартизованную процедуру в деталях невозможно; а если она и определена, то придерживаться ее, по всей вероятности, не будут из-за практических трудностей. Точно так же при сложной проблематике обследования в ходе опроса могут потребоваться уточняющие вопросы, а следовательно, гибкость и инициатива со стороны интервьюера. Вообще говоря, чем квалифицированнее, опытнее и заинтересованнее интервьюеры, тем больше доводов в пользу гибкости в выборе способа и формы подачи вопросов в ходе опроса.

При построении анкеты требуется сделать правильный выбор между весьма структурированными процедурами опроса и в значительной степени неструктурированным исследовательским подходом. В этой связи следует отметить несколько общих, но важных моментов. Во-первых, в большинстве случаев не желателен ни один из двух крайних подходов. Во-вторых, делая выбор, следует учитывать цели и условия проведения обследования; нельзя догматически отдавать безусловное предпочтение одному из них, как это иногда делают (см. ниже). В-третьих, структурирование анкеты связано с несколькими аспектами, причем некоторые могут быть стандартизованы в большей степени, чем другие. Эти аспекты таковы:

- 1) правила увязки и обобщения информации в анкете с целью обеспечения внутренней непротиворечивости полученных данных;
- 2) пояснения и указания для интервьюеров;
- 3) формулировка вопросов и способ их подачи респондентам;
- 4) структурирование вариантов ответов, т.е. форма регистрации полученной информации;
- 5) деление анкеты на разделы;
- 6) порядок вопросов и указания по переходу к следующему вопросу.

Эти аспекты поочередно рассматриваются в последующих параграфах.

Правила проверки непротиворечивости

В анкете следует предусмотреть возможность проверки внутренней непротиворечивости и/или достоверности информации, полученной интервьюером как в ходе опроса, так и при последующем редактировании в конторе. Сознательное введение избыточных пунктов может позволить осуществить контроль на разных этапах опроса. Анкета может включать дополнительные пункты о респонденте, условиях опроса, наличии третьей стороны во время его проведения и оценке интервьюером полученной информации.

В большинстве обследований уточняют процедуры опроса, выбор респондента, правила отмены и т.д. Когда такие детали учтены в каждой анкете, обеспечить соблюдение установленных процедур становится легче.

Инструкции для интервьюеров

Анкета представляет собой средство, позволяющее довести содержание обследования как до интервьюера, так и до респондента. Но важно четко различать этих лиц. Имеются примеры неудовлетворительного построения анкеты, когда это различие четко не было учтено. Иногда разработчик анкеты предусматривает детальные инструкции и правила для интервьюера, не уточняя, следует ли и в какой форме сообщать эти инструкции респонденту. Это легко может привести к утрате контроля над опросом и неодинаковому качеству работы.

По сути дела, вопросы можно формулировать двояким образом. Одна форма адресована преимущественно *интервьюеру*, для которого предусмотрены более или менее подробные инструкции для того, чтобы он хорошо понял смысл вопроса и соответствующие процедуры опроса; то, как необходимая информация сообщается респонденту, при этом не стандартизовано и не структурировано, а оставляется на усмотрение интервьюера. Другой вариант — адресовать анкету преимущественно *респонденту*, и обычно интервьюер должен придерживаться содержащихся в ней формулировок. Каждый из двух вариантов по-своему структурирует процесс опроса.

Вот пример из анкеты, рассчитанной в основном на интервьюера, который решает, какая еще требуется информация и как ее получить.

В. Характер основной активности индивида
Отрасль (см. перечень кодов)

Если код отрасли неясен, получите наименование и адрес предприятия. Если работает на собственном предприятии, опишите характер деятельности.

Ниже приведен пример вопроса, адресованного респондентам и взятого из одного недавнего обследования. Это пример неудачного подхода, так как неясно, нужно ли фразу в скобках зачитывать респонденту (а тогда бессмысленно заключать ее в скобки) или же она предназначена в качестве пояснения интервьюеру. В последнем случае неясно, как эта дополнительная информация, существенная для уяснения понятия “работа, приносящая доход”, будет доведена до респондента.

В. Выполняли ли вы какую-либо работу, приносящую доход (заработную плату, прибыль, семейный доход) в течение последних 12 месяцев?

Независимо от стиля формулировки вопросов почти в любой анкете содержатся хотя бы некоторые указания и пояснения для интервьюеров. Наличие таких указаний играет важную роль в деле структурирования опроса и обеспечения единых стандартов. Важные указания будут, по всей вероятности, проигнорированы, если они содержатся только в руководстве для интервьюеров, а не в самой анкете. Важный момент — это степень детализации указаний для интервьюеров в анкете. С одной стороны, важно включить ясные и подробные указания по всем существенным и сложным пунктам; например, интервьюер должен знать, кого опрашивать, как переходить к очередному вопросу и выдерживать последовательность опроса, где записывать ответы и, если необходимо, как уточнять ответы и получать дополнительную информацию. Основная анкета Обследования рождаемости в мире содержит удачный пример четких и подробных указаний для интервьюера (World Fertility Review, 1975). С другой стороны, не менее важно избегать перегрузки анкеты чрезмерными указаниями. После некоторой учебы и получения опыта большинство интервьюеров скоро приобретают навыки постановки многих вопросов, и часто для них становится излишним и нежелательным читать подробные указания при каждом опросе.

Формулирование и перевод вопросов

При формулировании вопросов важное соображение заключается в том, насколько полным должен быть вопрос и в какой мере интервьюер должен придерживаться этой формулировки в ходе опроса. Эту проблему следует решать на ранней стадии разработки анкеты, ибо от этого зависит размер анкеты (а следовательно, издержки

производства и требуемые полиграфические мощности), требования к переводу при многоязычных ситуациях, соответствующая форма обучения и метод опроса, а также качество получаемых данных.

Имеется много доводов в пользу полных вопросов. Обеспечив интервьюеров полными формулировками, можно добиться единообразного применения процедур опроса. Разумеется, некоторые вопросы таковы, что необходимо настаивать на том, чтобы интервьюеры задавали их строго в предложенной в анкете редакции, как бы квалифицированы и опытны они ни были. Сказанное относится не только к вопросам оценочного и субъективного характера. Даже применительно к вопросам так называемого фактического характера точные формулировки могут иметь важное значение, когда речь идет об отчетном периоде, искомой информации и о том, что следует включить в ответ и исключить из него. Полные вопросы способствуют также плавному переходу от одной серии вопросов к другой и правильному определению контекста новой серии.

Несмотря на перечисленные преимущества, существуют ситуации, когда точные и детальные формулировки всех вопросов невозможны, необязательны и даже нежелательны. Примеры, можно сказать, очевидны: при простых вопросах полные и подробные формулировки мало что дают; сложные ситуации, когда интервьюер не может строго придерживаться формулировок из-за разнообразия условий проведения обследования языковых и других особенностей респондентов; чрезмерно высокие издержки производства громоздких и подробных анкет; возрастание сложности проведения опроса на основе анкеты, при которой один и тот же вопрос сформулирован по-разному применительно к различным категориям респондентов. В последнем случае анкета может получиться сложной и длинной, и интервьюеру трудно будет придерживаться анкеты, видеть связь между различными элементами информации и контролировать полноту и внутреннюю непротиворечивость ответов в ходе опроса.

Подход, при котором вопросы даны не в виде точных формулировок, которых надо придерживаться, а в виде перечня вариантов, из которых следует выбрать правильный, именуют *списочным*. Важно отметить, что варианты при списочном подходе не обязательно определены менее подробно, чем при полных формулировках вопросов. Список в этом смысле не является сокращенной анкетой. Сокращены лишь форма и формулировка вопросов. Часто это также означает, что излишне повторять вопросы в списке в тех случаях, когда различные категории респондентов требуют слегка измененных формулировок, как это приходится делать в анкетах с полными

формулировками. По названным причинам объем анкеты списочного типа может быть значительно меньшим, чем равноценная анкета с полными вопросами.

Списочный подход удобен, когда сложность и разнообразие условий опроса требуют гибкости в способе получения информации; когда сложность тематики обследования и ограниченные способности респондента дать нужную информацию делают необходимым частое обращение к уточняющим вопросам и их постановке в альтернативной форме с целью получения информации и когда персонал настолько опытен и обучен, что на него можно положиться в деле правильной постановки вопроса в различных ситуациях. Списочный подход особенно пригоден в ситуациях, когда искомая информация носит преимущественно количественный характер и не зависит от точных формулировок или же нет сложных переходов между вопросами и их легко представить в компактной форме перечня.

При выборе подходящей формы вопросов следует сопоставить преимущества полных вопросов — возможность большего контроля, стандартизация и четкость — с преимуществами списочного подхода, а именно, его краткостью и гибкостью. По сути дела, речь идет о том, в какой мере анкета должна направлять интервьюера и респондента в ходе опроса. Как отмечено выше, это зависит от характера вопросов и условий проведения обследования.

Изложение вопросов в сокращенной форме не препятствует подробному опросу; это означает лишь большую свободу. Часто списочный подход предпочитают скорее по соображениям удобства и экономии, а не во имя гибкости и разнообразия в работе интервьюеров. Точно так же цель подробного изложения вопросов в полной анкете может заключаться главным образом в том, чтобы показать интервьюеру, как их следует подавать в *большинстве* случаев, а не навязывать их строгую постановку во *всех* случаях. И действительно, при упрощенном списочном варианте опасность может заключаться в том, что разработчики сами могут не знать, как следует ставить вопросы, так как они не пытались их строго сформулировать при построении анкеты.

Во многих случаях может оказаться удобным сочетание полных вопросов, сформулированных в анкете, и более сжатых перечней или таблиц, где лишь перечислена требуемая информация. Например, ключевой вопрос о том, в чем состояла экономическая активность индивида в отчетном периоде, может требовать максимально тщательной и точной формулировки в анкете. Точно так же вопросы о поиске работы должны быть правильно поняты таким образом,

что они включали как поиск работы по найму, так и попытки создания собственного предприятия. И наоборот, лишь немного может быть упущено, если ряд других вопросов сформулировать довольно кратко (например, “продолжительность отсутствия на работе”).

Опыт показывает, что обычно нет смысла в анкете приводить длинные вопросы с подробными описаниями и оговорками. В большинстве случаев интервьюеры считают такие вопросы слишком громоздкими: задавая их, они попросту игнорируют сложные оговорки. Если детали важны, лучше разбить вопрос на несколько более простых, а если они не столь важны, лучше их вообще игнорировать.

Когда полные и подробные вопросы задают в условиях многоязычной среды, важно перевести их на основные языки данной местности. Многие страны многоязычны или по меньшей мере в них говорят на нескольких основных диалектах. Формальный официальный язык может существенно отличаться от вариантов и языков, используемых обычно в стране, и это важное обстоятельство при проведении общенационального обследования. Особенно сложная ситуация имеет место во многих африканских странах. Очевидно, что проблему перевода нельзя игнорировать, так как она серьезно отражается на затратах и качестве полученной информации. Проблема языка при обследованиях рабочей силы изучена недостаточно, но можно сослаться на ценные исследования и опыт в рамках Обследования рождаемости в мире (см., например, Скотт и др., 1988).

Построение вариантов ответов

Открытые или закрытые вопросы

Не менее важный аспект разработки анкеты — это определение видов и диапазона ожидаемых ответов и их доведение до интервьюера и респондента. Двумя крайними случаями являются полностью “открытые” (допускающие множество ответов) и совершенно “закрытые” вопросы.

В первом случае респондент и интервьюер располагают максимальной свободой варьирования в регистрации ответов. Считается, что интервьюер задает вопрос, не влияя при этом на форму и характер ответа. Респондент вправе выбрать любую форму ответа, а интервьюер его полностью регистрирует, опять-таки не оказывая никакого воздействия на характер ответа. В другом крайнем случае четко оговорены варианты ответов, из которых респондент выбирает какой-либо один.

Кратко охарактеризуем преимущества того и другого подхода. “Открытые” вопросы позволяют избежать смещений, возможных в случаях, когда респонденту подсказывают варианты ответов, как это происходит при “закрытой” форме. Сохраняется при этом и непосредственность, спонтанность ответов. Такие вопросы больше подходят для уточнения данных. Открытые вопросы позволяют “срезать углы” (но обычно ценой ухудшения качества полученной информации): один более общий открытый вопрос способен заменить длинный ряд закрытых вопросов. На практике основаниями для применения открытых вопросов служат: а) отсутствие информации, необходимой для предварительного кодирования всех возможных и имеющих отношение к делу ответов; б) наличие слишком большого числа вариантов ответа, что не позволяет их заранее закодировать (например, занятия и отрасли). Но открытые вопросы обычно создают больше проблем при регистрации, кодировании и обработке, чем закрытые.

Содержательные и аналитические цели обследования часто предполагают получение определенных ответов, а это лучше достигается с помощью закрытой формы вопросов. Возможные преимущества таких вопросов включают: точное определение контекста и содержания вопроса; более удобную форму получения и регистрации ответов; как правило, лучшее регулирование и стандартизацию процесса опроса; меньший разброс в ответах; и прежде всего — упрощение задачи кодирования и обработки данных. Главное требование при использовании закрытых вопросов состоит в том, что разработчик анкеты должен быть способен выявить наиболее частые ответы, сгруппировать их в приемлемое число категорий без того, чтобы исказить фактически полученные ответы. Характер и число категорий ответов следует выбирать в зависимости от ожидаемого их разброса, целей вопроса и последующего анализа.

Важно отметить, что на практике существует целый континуум или градация структур между двумя крайними полюсами вопросов. Во-первых, многие вопросы *отчасти* открыты. Это означает, что большинство ответов заранее определены и закодированы, и включена открытая категория ответов (например, “другие, укажите...”) помимо тех, что были заранее закодированы.

На этой стадии будут полезны некоторые примеры. В большинстве обследований рабочей силы большая часть вопросов полностью закрытые; простейшим и наиболее распространенным примером этого служат вопросы, на которые единственно возможные ответы: “да” или “нет” (например, вопрос 21 на рис. 10). Наиболее распространенным примером открытого вопроса при обследованиях рабочей

21. РАБОТАЛИ ЛИ ВЫ БОЛЬШЕ, ЧЕМ НА ОДНОЙ РАБОТЕ,
НА ПРОШЛОЙ НЕДЕЛЕ?

Да (*переходите к B24*)

Нет

1

2

22. КАКОГО РОДА РАБОТУ ВЫ ВЫПОЛНЯЛИ?

.....
.....
.....

--	--	--	--

51. БЫЛА ЛИ ЭТА РАБОТА:

ЗА ЗАРАБОТНУЮ ПЛАТУ ИЛИ ЖАЛОВАНИЕ

(*переходите к B55*)

НА СОБСТВЕННОМ ПРЕДПРИЯТИИ

С НАЕМНЫМИ РАБОТНИКАМИ

БЕЗ НАЕМНЫХ РАБОТНИКОВ

БЕЗ ОПЛАТЫ НА СЕМЕЙНОМ ПРЕДПРИЯТИИ

(*переходите к B53*)

КАКОВА СИСТЕМА ВОЗНАГРАЖДЕНИЯ?

Оплата натурой (*переходите к B55*)

Неоплачиваемый добровольный труд

(*переходите к B68*)

1

2

3

4

5

6

53. ПРИЧИНЫ ОТСУТСТВИЯ НА РАБОТЕ НА ПРОШЛОЙ
НЕДЕЛЕ

Болезнь или травма работника

Отпуск/личные причины

Отсутствовала работа

Плохая погода/поломка

Забастовка/локаут

Прочие

1

2

3

4

5

6

Источник: Бюро статистики Австралии: Обследование рабочей силы, Австралия.
Канберра, 1985. Каталог №6203.0

Рис. 10. Примеры уточнения вариантов ответов

силы являются вопросы о занятии и отрасли (например, вопрос 22 на рис. 10). Вопросы о причинах временного отсутствия на работе, отказа от поиска ее и т.д. обычно отчасти закрыты: наиболее распространенные ответы перечислены, но имеется открытая категория ("прочие, укажите...") для случаев, не охваченных готовыми ответами (например, вопрос 53 на рис. 10). Предварительное определение вариантов ответов преследует несколько целей. Оно сокращает затраты труда на регистрацию и кодирование ответов; способствует получению ответов, представляющих особый интерес; помогает правильному течению опроса, что может зависеть от полученных ответов (например, вопрос 51 на рис. 10); способствует получению более сопоставимой информации в ходе нескольких раундов обследования.

Во-вторых, можно сформулировать вопросы, которые были бы полностью или частично закрыты для интервьюера (т.е. с точки зрения регистрации ответов), но были бы, по сути, открытыми для респондента (т.е. с точки зрения ответа). Респонденту задают с виду открытый вопрос (без определения вариантов ответа), тогда как интервьюер располагает заранее закодированным перечнем ответов, куда он заносит все или большинство полученных ответов. Основная цель такой процедуры — облегчить регистрацию и кодирование ответов, а не подсказка их респонденту. Другая возможность — вопросы могут различаться по диапазону "допустимых" ответов, явно или неявно сообщаемых респонденту. На одном полюсе варианты ответов могут быть ясны из самого вопроса (например, простые да — нет) или точно определены в формулировке вопроса. И напротив, в вопросе могут быть упомянуты лишь некоторые из наиболее распространенных вариантов ответа, скажем, в виде примеров, но не исключающие другие возможные ответы. Из других же вопросов возможные варианты ответов могут вообще не вытекать. Такие варианты позволяют в разной степени регулировать процесс получения ответов и получать различные ответы на один и тот же вопрос.

Рис. 10 построен следующим образом: прописные буквы используют для вопросов и вариантов ответов, которые интервьюер обязан полностью зачитать; строчные буквы предназначены для кодирования ответов респондентов, а курсив использован для указаний интервьюеру. Следует обратить внимание на различие между вопросами 51 и 53 на этом рисунке. Оба вопроса полностью сформулированы (об этом говорят прописные буквы). В вопросе 53 интервьюер не зачитывает респонденту варианты ответов, а регистрирует их в соответствии с заранее определенными (закрытыми) вариан-

тами (насколько это возможно). Следовательно, вопрос является закрытым (за исключением последнего варианта “прочие”) для интервьюера, но фактически открытым для респондента. В отличие от этого в вопросе 51 интервьюер обязан зачитать все варианты ответов по одному, начиная сверху, как часть вопроса¹. Таким образом, вопрос 51, по сути дела, является закрытым для интервьюера и респондента. Тем не менее он свидетельствует о наличии более разнообразных подходов: последний вариант, относящийся к вознаграждению, носит уточняющий характер, и ответы на этот вопрос являются закрытыми для интервьюера, но не для респондента.

В крупномасштабных обследованиях полностью или частично закрытые вопросы явно предпочтительнее, нежели открытые вопросы, в большинстве разделов анкеты. Открытые вопросы вводят лишь в том случае, когда на то имеются очевидные основания, например, когда закрытая форма безусловно не подходит, отсутствует информация для предварительного кодирования ответов или же, в ограниченных случаях, когда открытая форма сулит значительные удобства и краткость. Открытые вопросы чаще можно использовать в небольших предварительных обследованиях, а также при разработке и опробовании анкет обследования. Полные варианты ответов могут быть полезными для улучшения формулировки и формы вопросов и выявления вариантов ответов для “закрытия” вопросов при их последующем использовании в ходе основного обследования.

Вопросы, допускающие один или несколько ответов

Еще один важный аспект относится к различению вопросов, допускающих один или много ответов. В обследованиях рабочей силы, как и во многих других видах обследований, большинство вопросов допускает лишь один ответ. Имеются, однако, вопросы, которые способны дать более богатую и аналитическую более полезную информацию, если допустить множественность ответов. Кроме того, некоторые вопросы таковы, что настаивать на единственном ответе было бы неверным и трудно осуществимым подходом. Примерами, когда множественность ответов может оказаться целесообразной, могут служить такие вопросы, как шаги по поиску

¹ Кстати говоря, указаниями в этом случае предусмотрено, что чтение вопроса можно фактически прекратить, как только получен положительный ответ. Это указание в анкете не приведено, но можно предположить, что оно содержится в руководстве для интервьюера или сообщено при инструктаже. Если интервьюеры неопытны, указание полезно поместить в самой анкете.

работы, причины неготовности приступить к ней или отказа от работы и т.д.

Но допустив множественность ответов, можно столкнуться с проблемами. Проблемы, как правило, возникают на стадиях опроса, кодирования и обработки данных. Результаты опроса при этом труднее анализировать и представлять. В целом невозможно отдать предпочтение одному из ответов или же для этого следует прибегать к дополнительным вопросам или произвольным допущениям. Поэтому при крупномасштабных обследованиях, как правило, рекомендуется избегать вопросов, допускающих множественность ответов, если только это не обеспечивает явные аналитические преимущества. Вместо того, чтобы просто настаивать на единственном ответе, это можно сделать, к примеру, спросив о “главной” причине или о “последнем” виде активности. Еще один подход заключается в том, чтобы перечислить варианты ответов в какой-то осмысленной системе приоритетов и считать ответом первый положительный ответ (как это сделано в вопросе 51 на рис. 10). Можно также превратить каждый вариант ответа в отдельный вопрос типа “да — нет”. Это, разумеется, удлинит опрос, но способно улучшить качество полученной информации. Перечень видов активности, упомянутый в гл. 2, может служить примером такого решения.

Членение анкеты на разделы

Большинство анкет любого размера удобно разделить на части в зависимости от: источника информации (кто ее представляет); единицы, к которой она относится; объекта опроса и, если это существенно, метода получения информации.

В большинстве обследований рабочей силы информацию делят на ту, которую получают на уровне домашнего хозяйства (анкета для домашнего хозяйства), и ту, которая относится к отдельным его членам (индивидуальная анкета). Можно использовать отдельную анкету для каждого лица. В рамках каждой анкеты можно предусмотреть разделы по основным темам (основные данные, активность в данный момент, обычная активность и т.д.), что облегчает инструктаж, сбор и обработку данных.

Особое значение имеет вводный раздел анкеты. При любом обследовании необходима информация: сведения об организации, которая проводит обследование; характеристика обследования или его раунда; адрес и расположение единицы, предоставляющей ответы; характер выборки; цифровые данные, позволяющие проконтролировать поток документов, обработку данных, связь между обследо-

ваниями или отдельными раундами, включая увязку индивидуальных данных с показателями домашнего хозяйства или семьи; аспекты опроса, например, дата, результат, детали уточнений в анкете, фамилии интервьюеров, а также контролеров и другого персонала; сведения о контроле качества, если он проводится. Всю перечисленную информацию целесообразно поместить во вводном разделе анкеты, как правило, на титульном листе. В этом же разделе могут содержаться указания для интервьюера и ряд вопросов, призванных облегчить начало опроса. Наконец, там же можно поместить основные показатели и выводы на основе полученной при опросе информации, что позволит ускорить табулирование важнейших результатов в предварительном виде.

Располагая разделы анкеты, следует принимать во внимание ряд факторов. Первой должна идти информация для определения характера и контроля опроса. Обычно опрос следует начинать с относительно простых вопросов описательного характера. Сложные аспекты не должны появляться слишком рано. Но, с учетом сказанного, вопросы, имеющие принципиальное значение с точки зрения основных задач обследования, предпочтительно не откладывать. Важно избежать такой последовательности разделов, при которой ошибочное толкование определенного термина или фразы затем невольно (или ошибочно) переносится на последующий раздел. Необходимость задавать одни вопросы раньше других может также диктоваться логикой ситуации. Например, некоторые вопросы необходимы для выявления респондентов, от которых потребуются ответы на последующие вопросы или на различные серии вопросов.

Чередование вопросов и переходы

Внутри каждого раздела вопросы следует располагать в удобном и логическом порядке. Основная задача состоит в обследовании главного опроса без неуклюжих и нелогичных скачков. При определении последовательности вопросов в анкете необходимо исходить из логики *респондента*, а не из логики аналитика. Важно, чтобы респондент придерживался определенной системы, и облегчить ему процесс вспоминания. Респондент должен хорошо сознавать, когда меняются контекст, предмет опроса, отчетный период, единица обследования и т.д. в ходе опроса. Однако вопросы следует расположить таким образом, чтобы свести эти изменения к минимуму. Если обсуждается какая-либо тема, все относящиеся к ней вопросы должны быть сведены воедино до перехода к другой теме.

Вопросы должны, как правило, идти в хронологическом порядке. Удачная последовательность вопросов часто помогает респонденту предвосхитить следующие вопросы, что облегчает проведение опроса.

В анкете необходимо предусмотреть однозначные указания о ходе опроса, о том, какие вопросы следует задавать тому или иному респонденту: их именуют “указаниями о переходе”. Их цель — избежать постановки вопросов, не имеющих отношения к данному респонденту; облегчить работу интервьюера, с тем чтобы он ставил вопросы, уместные в данной ситуации; ускорить обработку данных, обеспечив ход опроса на основе заранее установленных правил и с учетом уже полученных ответов. В сложных анкетах, рассчитанных на различные категории респондентов, переходы могут оказаться довольно замысловатыми. Особенно часто это имеет место при применении анкет с полными, законченными вопросами, которые, даже будучи одинаковыми по смыслу, могут быть сформулированы по-разному с учетом особенностей различных категорий респондентов. Следует всячески стремиться к тому, чтобы переходы не были слишком сложными для интервьюера и он мог осуществлять их безошибочно. Опрос никогда не должен возвращаться к предшествующей части анкеты, хотя иногда ссылки на нее возможны. Делать это следует исключительно для целей сравнения, контроля или переноса информации. Иногда переходы можно упростить с помощью так называемых “фильтров”; под ними понимаются те места в анкете, где суммируется информация с одного или нескольких предшествующих вопросов с целью обеспечить более удобные переходы. Повторив ряд вопросов, чего можно было бы избежать при иной структуре анкеты, иногда удастся упростить переходы. Вообще говоря, опыт показывает, что при должном построении анкеты во многих случаях удастся избежать переходов, которые затрудняют работу и ведут к ошибкам со стороны интервьюера, при условии, что разработчик анкеты не очень скупится на размеры анкеты и предусматривает отдельные серии вопросов для различных категорий респондентов. Затрудняющие работу переходы особенно часты тогда, когда разработчик чрезмерно озабочен экономией места или же просто неудачно устанавливает очередность вопросов. Графическое представление анкеты (сетевая, или поточная диаграмма рассмотрена выше в настоящей главе) может оказаться весьма полезным инструментом при разработке сложной анкеты и определении должной последовательности вопросов и переходов между ними.

5. РЕАЛИЗАЦИЯ АНКЕТЫ

Проект и проверка

Разработка хорошей анкеты — это техническая задача, требующая высокой квалификации, здравого смысла, а также знания предмета и условий обследования. Некоторые технические принципы и удачные примеры построения анкеты были рассмотрены выше.

Как только анкета разработана на основе этих принципов, ее следует опробовать на практике. Может потребоваться не одна проверка, прежде чем будет получен вариант, пригодный для широко-масштабного обследования.

Важность тщательного анализа и проверки анкеты невозможно переоценить. В практике обследований нередки случаи, когда с виду крайне незначительные и чисто случайные ошибки, оставаясь незамеченными в анкете, приводили к серьезным неудобствам для интервьюеров и обработчиков данных и к ненужным изъясам в обследовании.

Проверка анкеты — это обычно итеративный процесс, в ходе которого вариант, улучшенный в результате предшествующего контроля, подвергается дальнейшему опробованию и оценке до получения удовлетворительного результата. В процессе анализа и проверки проекта анкеты можно выделить следующие шаги (United Nations, 1985, p. 232—244).

1) *Технический анализ* бригадой, проводящей обследование, включая анализ ожидаемого выхода совместно с пользователями. Тщательный, слово за словом, анализ проектов анкеты — это необходимый шаг: проверка в деле (“полевые испытания”) не должна заслонить важность анализа в конторе. Желательно, чтобы в анализе участвовал широкий круг специалистов, контрольный персонал и пользователи на различных уровнях. Анализ должен затрагивать общие аспекты обследования, а также техническую сторону анкеты. Часто бывает полезно организовать анализ в виде совещания. Его участников следует снабдить контрольным листом, что обеспечит им работу, ход совещания необходимо запротоколировать, равно как отразить все замечания и принятые решения. Как отмечалось в разделе 1, следует тщательно проверить, насколько учтены эти замечания в проекте анкеты, эту работу следует поручить небольшой бригаде квалифицированных работников.

2) *Предварительное опробование на местах*. Его цель — оценить, как воспринимается анкета, и выявить возможные проблемы при контактах с респондентом. Предварительное обследование призва-

но, как правило, ответить на следующие вопросы. Хотят ли и могут ли различные категории респондентов ответить на задаваемые вопросы? Протекает ли опрос гладко, и легко ли интервьюеру пользоваться анкетой? Достаточно ли места для регистрации ответов, и ясны ли указания о переходах между вопросами? Сколько времени в среднем длится опрос? Возникают ли вопросы особо трудные или ведущие к ошибкам?

Предварительное опробование может быть сравнительно небольшим по своему масштабу, скажем, 100 — 200 опросов на основе сознательно отобранной выборки, включающей все остальные категории респондентов в городе и деревне. В отдельных случаях, например, для составления вариантов ответов и кодов, особенно когда следует сопоставить несколько подходов, выборка может быть значительно большего объема. Но в целом упор в ходе предварительного опробования необходимо сделать на удобство и практичность, тесный контроль и наблюдение, быструю обратную связь и оперативное использование результатов, а также разрешение возникших конкретных проблем.

С учетом сложности обслуживания и прежнего опыта может оказаться необходимым не одно предварительное опробование. Если первая проверка вылилась в значительные исправления анкеты, желательно провести повторное опробование внесенных изменений.

3) *Пробное ("пилотажное") обследование*. Если основная цель предварительного апробирования анкеты — выявление и снятие конкретных проблем, то пробное обследование — это последняя "генеральная репетиция" всей процедуры обследования до проведения полномасштабной работы. При этом следует по возможности воспроизвести реальные условия обследования. Оно должно охватывать более обширную и представительную выборку. Персонал, организация опроса, наблюдения, контроль качества и т.д. — все это должно быть таким же, как и при полном обследовании. В идеальном случае подлежат охвату все этапы, включая сбор данных, их кодирование, обработку и табулирование.

4) *Проверка альтернативных подходов*. Иногда имеющиеся опыт и знания недостаточны для того, чтобы сделать выбор между альтернативными подходами, и может возникнуть потребность в формальном, научно обоснованном сопоставлении. Планирование и проведение экспериментов, которые могут привести к однозначным решениям и выбору, — эту сугубо специальная задача, требующая знания статистической методологии. Эти эксперименты могут требовать значительных средств и времени. Различные альтер-

нативы следует опробовать в реальных условиях, а объем выборок должен быть достаточно большим, чтобы можно было с достаточной точностью оценить различия в затратах и результатах. Чтобы добиться успеха, необходимо четко сформулировать гипотезы и критерии выбора между альтернативами; следует определить условия эксперимента и проследить, чтобы они выполнялись на практике. Многие проверки не дают полезных результатов, так как их цели не определены достаточно четко или же слишком амбициозны, а условия проведения недостаточно контролируются. Вопросы экспериментальных проверок и обследований рассмотрены в работе (Jabine, 1982).

Компоновка материала и печать

При крупномасштабном производстве анкет для большого обследования важно уделить внимание техническим аспектам анкеты — размеру, форме, полиграфическому исполнению, кодированию и другим требованиям обработки данных, а также тиражированию в достаточных количествах. Необходимо также подготовить различные вспомогательные материалы и документы до начала работы по сбору данных. В каком-то смысле указанные аспекты разработки анкеты меньше зависят от конкретного содержания обследования. Они требуют чертежного искусства, умения компоновать материал и чувства формы, знания процедур обработки данных и знакомства с техникой воспроизведения документов.

Вопросы воспроизведения анкет и смежные проблемы рассмотрены в работе ООН (1985, с. 190—224), где также содержатся примеры из ряда стран. Примеры компоновки материалов в анкетах обследования проведены также в приложениях к настоящей книге. Ниже обобщены некоторые наиболее важные моменты физического построения анкет.

Как только анкета опробована и уточнено ее содержание, необходимо тщательно отработать ее физическую форму и компоновку до отправки в печать и на размножение. Неудачное расположение материала является источником случайных ошибок при сборе, регистрации, кодировании и обработке данных. При окончательной компоновке следует учитывать ряд моментов, в том числе:

- 1) *Идентификация анкеты.* Все анкеты должны иметь место для идентификационных номеров, а также адреса и другой требующейся описательной информации. При продолжающихся крупномасштабных обследованиях особенно важно разработать хорошую систему нумерации, которая обеспечила бы увязку между различными

ми раундами, подвыборками и видами анкет в рамках одного обследования.

2) *Нумерация вопросов.* Система призвана обеспечить безошибочную идентификацию, ясность и удобство при сборе и обработке данных.

3) *Место для регистрации ответов.* Следует четко указать форму и единицу, в которых записывается ответ. Пространство должно быть достаточным для ответа и не должно быть путаницы в отношении того, к какому вопросу данный ответ относится.

4) *Указания для интервьюера.* Весьма желательно, чтобы по меньшей мере в наиболее важных случаях в анкете содержались указания для интервьюера о том, кто есть респондент и к кому относится информация; о переходах и фильтрах, регулирующих ход опроса; о форме и единицах регистрации ответов на каждый вопрос; и, выборочно, об уточнении ответов на отдельные вопросы или трудные разделы. К примеру, следует четко различать вопросы, допускающие единственный или множественные ответы. Вообще говоря, указания для интервьюера нужны в каждом случае, когда требуется какое-либо особое действие в ходе опроса.

5) *Полиграфическое разнообразие.* Там, где оборудование позволяет, можно значительно улучшить восприятие и внешний вид анкеты путем подчеркиваний, применения жирного шрифта, различных кеглей, цветов и т.д., что облегчает различение частей анкеты и видов информации, сообщаемой интервьюеру и респонденту. См., например, рис. 10, рассмотренный выше.

6) *Требования, связанные с обработкой данных.* Окончательный вариант анкеты должен включать детали, относящиеся к дальнейшей обработке собранных данных. Речь идет о единицах измерения для каждого случая, выборке системы кодирования и кодах, вносимых в анкету, инструкциях о вводе данных (см. также United Nations, 1982).

7) *Физический размер и форма анкеты.* Этим аспектам следует уделить внимание по соображениям затрат, а также удобства опроса и обработки данных. При комплексных обследованиях громоздкие анкеты могут представлять собой проблему. Иногда экономичнее и/или удобнее разделить анкету на самостоятельные части, особенно если они относятся к различным респондентам или выборкам. Но при этом надо обеспечить, чтобы эти части можно было впоследствии объединить или увязать.

Говоря о печатании анкет, необходимо отметить два момента. Во-первых, следует учитывать количество требуемых экземпляров. Печатать слишком много расточительно, однако нехватка их может

быть пагубной или по меньшей мере повлечь за собой неудобства и затраты. Помимо экземпляров, необходимых для опроса, следует напечатать некоторое количество анкет для обучения интервьюеров, а также с учетом возможной их порчи, ошибок при распределении тиража, неопределенности в объеме выборки, отправки заинтересованным лицам и организациям.

Во-вторых, экземпляр, направленный в типографию, и гранки, полученные из типографии, необходимо тщательно вычитать, чтобы выявить типографские ошибки, а каждую партию следует проверить выборочно с точки зрения качества исполнения работы. Во многих случаях ошибки или дефекты на стадии полиграфического исполнения портили работу, тщательно проделанную на более ранних этапах.

Литература

- Australia. 1985. *Labour Force Survey, Australia*. Canberra, Australian Bureau of Statistics, Catalogue No. 6203.0.
- Casley, D.; Kumar, K. 1988. *The collection, analysis and use of monitoring and evaluation data*. Baltimore, Johns Hopkins University Press.
- International Labour Office (ILO). 1983. Thirteenth International Conference of Labour Statisticians: *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*.
- 1983—84. ILO methodological survey on the measurement of employment, unemployment, underemployment and income in Kerala, India. Unpublished document, Geneva, ILO Bureau of Statistics.
- 1986. “La encuesta metodológica de la OIT para la medición del empleo, del desempleo y del subempleo en Costa Rica”, by R. M. Trigueros (with summaries in English and French), in *Bulletin of Labour Statistics* (Geneva) 1986—1, pp. ix-xx.
- Jabine, T. 1982. *Guide-lines and recommendations for experimental activities and pilot surveys*. PIDEH Working Paper No. 1 (document 7679a). Washington, DC, Inter-American Statistical Institute.
- 1983. “Teaching questionnaire design”, in *Statistical Review*, 5, pp. 157—166.

Scott, C. et al. 1988. "Verbatim questionnaire versus field translation or schedules: An experimental study", in *International Statistical Review*, 56 (3), pp. 259—278.

United Nations. 1982. *Survey data processing*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.

— 1985. *Development and design of survey questionnaires*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.

World Fertility Survey. 1975. *Core questionnaires*, WFS Basic Documentation, No. 1. The Hague, International Statistical Institute.

— 1976. *Editing and coding manual*. WFS Basic Documentation No. 7. The Hague, International Statistical Institute.

Глава 13. ПРОВЕДЕНИЕ ОБСЛЕДОВАНИЯ И ОБРАБОТКА ДАННЫХ

1. ВВЕДЕНИЕ

В этой главе рассмотрены некоторые практические вопросы проведения обследования рабочей силы. С учетом этапов планирования и проектирования, обсужденных в главах 10 — 12, основные работы в ходе обследования включают: 1) набор и обучение персонала; 2) проверку инструментов и процедур обследования; 3) подготовку основы выборки и отбор элементов; 4) сбор данных на местах; 5) подготовку и обработку данных; 6) оценку данных; 7) табулирование, составление отчета и рассылку.

Каждая из перечисленных работ может включать несколько различных шагов. На практике работы могут пересекаться и различным образом переплетаться. Эти работы поочередно рассмотрены ниже, за исключением оценки данных, которая подробно освещается в гл. 14.

Обсуждая различные работы, следует подчеркнуть, что не существует единой модели проведения общенационального обследования рабочей силы. Страны существенно различаются по условиям, ресурсам и организации статистических работ; обследования же различаются по своим целям, сложности, структуре и масштабам. Тем не менее возможно и полезно выявить основные варианты и факторы, определяющие их выбор, а также рекомендовать некоторые общие удачные подходы. Рекомендации по практическим аспектам проведения обследования содержатся во многих руководствах и статьях. Особенно полезна работа ООН “Справочник по обследованиям домашних хозяйств” (1984 г.), где содержится также обширная библиография по этому вопросу. Настоятельно рекомендуем серию основных документов “Обзоры рождаемости в мире”. Важно различать отдельные этапы работ, так как каждый из них требует особого подхода и ресурсов в рамках планирования и реализации проекта. К сожалению, это существенное требование не всегда соблюдают на практике. Например, часто собирают большое количество сложных данных, не предусмотрев выделения необходимых ресурсов для обеспечения их быстрой и точной обработки и анализа. Есть примеры организации обследований, когда внимание акцентируют только на сборе и табулировании данных, игнорируя другие важные аспекты — обучение персонала, предварительное опробование, контроль качества, оценку и рассылку данных. Многие обследования не достигают цели не потому, что общих ресурсов

недостаточно, а главным образом из-за их неправильного распределения. Последнее зависит, разумеется, от характера обследования. Так, при разовых обследованиях на основе особых методов и новых процедур требуется уделять больше внимания набору и обучению персонала, планированию обследования и проверке анкеты, чем при проведении обследований рабочей силы устоявшимися методами. Но и в последнем случае было бы ошибкой игнорировать эти аспекты, особенно если внесены изменения в планы и процедуры. Хотя перечисленные выше работы даны в относительно хронологической последовательности, на практике они могут пересекаться и различным образом переплетаться, и это неизбежно при непрерывном обследовании рабочей силы, когда несколько работ проводят одновременно. Кроме того, более тесная интеграция разных работ становится возможной благодаря последним достижениям в использовании компьютеров для проведения обследований. Традиционный подход к сбору и обработке данных при крупномасштабных обследованиях предполагает значительное число самостоятельных этапов: анкеты проектируют и komponуют вручную, затем их набирают и печатают; отбор выборки, назначение на работы и отчеты о результатах выборки также обычно выполняют вручную; анкеты, заполненные в ходе опроса, передают редакторам для контроля, а затем в центральный офис для дальнейшего ручного контроля и кодирования, причем на каждом этапе ведется строгая документация движения бумаг; документы, содержащие ошибки, возвращаются на места, а затем уже направляются в вычислительный отдел для ввода данных, машинного редактирования и т.д. Машинное редактирование само по себе связано с рядом шагов по проверке структуры и непротиворечивости данных, изменению их формы, добавлению отсутствующих данных и т.д. Программы редактирования дают массу ошибок, исправление которых требует обращения к анкетам. Часто требуется неоднократное повторение цикла редактирования, прежде чем данные становятся пригодными для табулирования. Суть традиционного (и все еще распространенного) подхода состоит в том, что одновременно в работе находится один этап, который заканчивают применительно ко всей выборке, прежде чем переходить к следующему этапу.

Достижения в области компьютерной технологии позволяют сочетать несколько этапов. Важным достижением является интерактивный ввод данных, который именуют "вводом данных с помощью компьютера" (CADI). Он связан с использованием компьютера для регулирования ввода и контроля данных по мере их поступления. Это позволяет объединить подготовку и ввод, а также различные

стадии редактирования данных в одну операцию. Еще более высокая степень интеграции достигается в тех случаях, когда интервьюеры используют компьютеры для регистрации ответов в ходе самого опроса (опрос с помощью компьютера, CADI). В стадии разработки находятся более общие системы проведения обследований с применением компьютеров, обеспечивающие дальнейшую интеграцию работ от определения формы анкеты, последовательности вопросов и хранения данных о выборке до ввода, редактирования, увязки данных и выдачи окончательных массивов и таблиц. Интерактивный ввод данных получает быстрое распространение, в том числе и во многих развивающихся странах. Сбор данных и более общие системы обследований на базе компьютеров не столь распространены, но и они применяются в более развитых странах, например, в Нидерландах, а в последнее время и в Великобритании.

2. НАБОР И УЧЕБА ПЕРСОНАЛА

Организация и персонал обследования

Любое обследование предполагает наличие персонала определенной квалификации. Можно назвать три основные категории работников:

специалисты, знающие предмет обследования, системы управления, выборочный метод, вопросы разработки анкет и обработки данных, а также статистики, знакомые с различными аспектами обследований;

менеджеры среднего уровня, в чьи функции входит контроль за проведением обследования;

оперативный персонал — интервьюеры, кодировщики и другие работники.

Приобретение квалификации — длительный процесс. Количество специалистов высокой квалификации в организации, проводящей обследование, не должно быть очень большим; важно лишь, чтобы организация имела доступ к лицам нужной квалификации, которых можно было бы привлекать для выполнения специальных работ. Наиболее важная категория специалистов, которые в идеальном случае должны быть заняты в организации на постоянной основе, это, пожалуй, статистики-специалисты по обследованиям. Это лица, знакомые со всеми аспектами обследования — от планирования и обучения до сбора, обработки и представления информации, даже если они не являются специалистами в какой-либо отдельной области; они должны уметь контактировать со специалистами и принимать разумные решения.

Многие статистические органы страдают, и нередко серьезно, из-за дефицита персонала среднего уровня, хорошо знающего практические аспекты переписей и способного организовать их проведение. Такие люди нужны для помощи, например, в обучении работников, проведении предварительного опробования анкет, подготовки документов для выборки, контроля за опросами, кодированием и редактированием. Эти люди должны быть заняты в организациях, регулярно проводящих крупные обследования, полное рабочее время. Обучению этих лиц с учетом их конкретных функций следует уделять первостепенное внимание.

Оперативный персонал обычно требуется в значительно больших количествах. С учетом обстоятельств возможны различные варианты комплектования штатов. (Эти вопросы рассмотрены в работе ООН, 1984, парагр. 5.3 — 5.17.) Основной выбор заключается между постоянным штатом и специально набранным персоналом. Возможны различные степени постоянства штата на основе заключения трудовых договоров на разные сроки. Каждый вариант имеет свои преимущества и недостатки, однако для постоянных обследований рабочей силы часто предпочтительнее постоянный или по меньшей мере регулярно занятый оперативный и конторский персонал. Постоянный штат обладает преимуществами в виде накопленных знаний и опыта и меньших затрат на обучение. Но этот вариант может оказаться неэффективным, если требуется гибкость в объеме и характере работ в различные периоды времени. Имеются случаи, когда организации не могут загрузить работой всех постоянных сотрудников.

Другая проблема комплектования штатов — нанимать людей на полное или неполное рабочее время. Выбор зависит от наличия кандидатов и характера работы. Как правило, при крупномасштабных регулярных обследованиях в развивающихся странах предпочитают занятых полное время, но часто им поручают несколько функций, а в обследованиях они заняты с перерывами и неполное рабочее время.

При наборе кадров и распределении работ во многих странах необходимо также принимать во внимание языковое и этническое разнообразие обследуемого населения. В определенных случаях необходимо (либо желательно или дешевле) набирать людей из числа местных жителей в выборочном районе. Это, разумеется, может уменьшить гибкость в организации работ и при изменении плана обследования.

Количество персонала зависит от таких факторов, как размер выборки, время, отведенное для завершения работ, и производительности. Надежные оценки последнего фактора (на основе про-

шлого опыта, предварительного обследования и т.д.) особенно важны при обследованиях рабочей силы с их жесткими требованиями к срокам, в частности, когда обследование разделено на короткие раунды и подраунды, каждый из которых следует завершить к определенному сроку. Определяя потребность в персонале, необходимо учитывать выбытие по разным причинам, особенно при непрерывных обследованиях. Поправки на выбытие будут, как правило, более значительными при использовании временных, а не постоянных работников; при сложных обследованиях, требующих особых усилий или происходящих в трудных условиях; при работе в одиночку, а не бригадами (когда работу можно распределить между ее членами при возникновении трудностей); в ситуациях, когда при распределении работ имеются ограничения языкового, этнического или географического характера. Перечисленные факторы предполагают также наем относительно большего числа контролеров. Как правило, один контролер должен приходиться на пять-семь интервьюеров (или кодировщиков); в более сложных ситуациях может потребоваться контролер на каждые два-три интервьюера.

Обучение

Обучение — ключ к качественным данным. Объем учебы зависит от подготовки и опыта персонала и сложности обследования. Специально нанятый или случайный персонал требует более основательного обучения, нежели работники, участвовавшие в предыдущих обследованиях. Но и при непрерывных обследованиях с постоянным персоналом периодически необходима переподготовка для проверки и обновления знаний. Обследования рабочей силы связаны со сложными понятиями, которые надо применять в разных условиях. Ни одно обследование не является простым повторением от раунда к раунду. Рутинный подход может привести к тому, что работники лишаются заинтересованности, усваивают плохие приемы, которые необходимо выявлять и исправлять.

Необходимо обратить особое внимание на учебу *работников среднего звена*. Она требует постоянных и продолжительных усилий. Часто приходится проводить учебу до того, как определены все детали плана и процедуры обследования. Как правило, желательно централизовать учебные программы, с тем чтобы добиться единообразия подхода, хотя это может увеличить затраты. Другая распространенная проблема — отсутствие старших руководителей для проведения учебы. Для того чтобы впоследствии обучать и контролировать работу оперативного персонала, работники среднего звена

должны хорошо понимать и принципы, и конкретные детали обследования. Учеба должна включать как теорию, так и практику, формальный инструктаж с неформальными выходами на места. Участие в предварительном обследовании, составлении списков и отборе выборки, а затем и в качестве помощников при обучении интервьюеров и кодировщиков и т.д. — таковы основные методы практической учебы персонала среднего звена. Кроме того, необходима более формальная учеба (скажем, в течение одной-трех недель в зависимости от обстоятельств), нацеленная на приобретение ими навыков управления людьми и контроля качества.

Оперативный персонал надо обучать конкретным деталям обследования. Отдельные курсы необходимы для обучения составлению списка единиц отбора, опросу, кодированию и редактированию (сверке) данных. Как правило, на это требуются две-четыре недели формального обучения каждой функции. Курсы для интервьюеров могут, к примеру, включать следующие темы: общая информация об обследовании, включая организационные и плановые аспекты; детальное разъяснение понятий и вопросов, включенных в обследование; указания на случай возникновения сложных ситуаций; техника опроса; процедуры проверки и корректирования полученной информации. А так как число учащихся обычно велико, необходимо уделить особое внимание условиям и организации учебного процесса. Эффективная учеба предполагает сочетание классных занятий с достаточной практической работой. Ученики-интервьюеры могут, например, сочетать классные занятия с самостоятельной учебой, присутствовать при учебных опросах, участвовать в имитациях опроса (когда ученики опрашивают друг друга под наблюдением инструктора), а также проводить опросы в реальных условиях. (Методы обучения хорошо обсуждены в World Fertility Survey, 1976). Важно оценить работу учащегося в процессе учебы. Эффективным способом в этом смысле является запись опроса на пленку.

В крупных и многоязычных странах обучение оперативного персонала, возможно, придется децентрализовать, и в этом случае следует позаботиться об единых стандартах путем всестороннего обучения учащихся в центральном пункте и максимально подробного описания рекомендуемых процедур проведения обследования.

Число лиц, обученных для каждого вида работ, должно превышать потребное их количество с учетом возможного выбытия. Опыт показывает, что необходимо набирать для учебы примерно на 20% больше лиц, чем это потребуется для проведения работ; когда же набирают лиц специально или на временной основе, разрыв, вероятно, должен быть еще больше.

При непрерывных обследованиях рабочей силы с использованием постоянного персонала могут оказаться достаточными периодические курсы повышения квалификации сравнительно небольшой продолжительности. Однако они не должны быть слишком краткими, и на них не следует смотреть как на рутинное дело. Чтобы они принесли пользу, необходимо тщательно оценить прошлый опыт, выявить конкретные проблемы и найти способы их решения. Следует тщательно рассмотреть на курсах все изменения в планах и процедурах обследования. Курсы повышения квалификации призваны также помочь оперативному персоналу довести до инструкторов проблемы, с которыми они столкнулись при обследовании. Эти курсы могут сыграть важную роль в деле улучшения климата и мотивации сотрудников.

3. ОПРОБОВАНИЕ ИНСТРУМЕНТОВ И ПРОЦЕДУР ОБСЛЕДОВАНИЯ

До проведения крупномасштабного обследования необходимо опробовать и оценить процедуры. Проверяемые аспекты и масштабы проверки зависят от конкретных требований и прошлого опыта. Если предпринимают новое обследование или ведущееся обследование существенно видоизменено, опробование должно проводиться в сравнительно крупном масштабе. Более крупные и представительные выборки обычно необходимы тогда, когда ставится цель оценить альтернативные планы и процедуры и сделать выбор между ними. С другой стороны, при установившемся обследовании рабочей силы, тщательном ведении и анализе документации предыдущих раундов может отпасть необходимость в отдельном предварительном опробовании, если не вносятся изменения в действующие планы и процедуры.

Опробование до полномасштабной реализации проекта преследует несколько целей. Как уже отмечалось в парагр. 5 гл. 12, большая часть предварительных проверок направлена на опробование анкеты в реальных условиях. С целью выявить проблемы в постановке вопросов часто достаточно ограничиться относительно небольшим количеством опросов (скажем, 100—300) на основе целенаправленной выборки. Однако выборка должна включать все важнейшие категории респондентов изучаемой совокупности. Но обследование такого небольшого масштаба обычно не дает полезной качественной информации о дисперсии, издержках и других параметрах, необходимых для улучшения плана обследования, и она крайне недостаточна, чтобы сделать выбор между альтернативными планами и процедурами. Для этой цели, как правило, необходимы проверки на основе более крупных и представительных выборок. В парагр. 5

гл. 12 отмечалось также, что в отличие от предварительного опробования с ограниченными и конкретными целями, *пробное* (пилотажное) обследование означает более общую “генеральную репетицию” всех рабочих аспектов обследования до его полномасштабного проведения. Эти пробные обследования следует проводить более реалистичным и представительным образом, нежели опробования, преследующие более ограниченные цели.

Предпринимая любое опробование, важно: а) определить и ограничить его цели; б) установить четкие критерии оценки результатов и выводов на их основе; в) выработать процедуры сбора и отражения требуемой информации; г) тщательно проанализировать полученную информацию. Эти основные принципы на практике соблюдают не всегда. Очень часто опробование превращается в простую формальность, оно слишком мало и ограничено по охвату, но преследует слишком широкие и неопределенные цели, и мало что дает с точки зрения анализа и использования полученной информации.

Проверку анкет и процедур на местах следует проводить задолго (скажем, за три и более месяцев) до начала основной работы по сбору информации, с тем чтобы было достаточно времени для анализа результатов и внесения необходимых изменений. Следует также отметить, что при введении в обследование новых и сложных аспектов может потребоваться серия предварительных испытаний.

Еще одно важное соображение при организации предварительных опробований — это их потенциальная роль в обучении, особенно контрольного персонала среднего звена, о чем говорилось в предыдущем параграфе. Многие организации пришли к выводу, что целесообразно, насколько это возможно, использовать будущих контролеров в качестве интервьюеров в ходе таких опробований. Сочетание учебы с опробованием — это явно экономичный прием. Кроме того, персонал, уже знакомый с работой статистической организации, нередко более надежен как в смысле правильного применения проверяемых процедур, так и в отношении обобщения полученного ими опыта.

4. ОСНОВА ВЫБОРКИ И ФОРМИРОВАНИЕ ВЫБОРКИ

Обновление основы выборки

В большинстве случаев основы выборки или главные выборки *территориальных* единиц имеются или могут быть получены из какого-либо внешнего по отношению к обследованию источника, например, последней переписи населения. Обычно такие основы разрабатывают и ведут для обслуживания ряда обследований. Соображения, относящиеся к основе выборки (или главной выборке),

подробно рассмотрено в парагр. 3 гл. 11. Здесь же нас интересует *дополнительная* работа, которая может потребоваться для получения окончательной выборки для определенного случая, например, одного раунда непрерывного обследования рабочей силы. Это предполагает один или несколько шагов следующего рода:

1) Картирование, группировка, дробление или другие изменения в некоторых территориальных единицах основы выборки.

2) Отбор территориальных единиц для выборки, как это требуется для обследования.

3) Возможно, последующие ступени отбора (каждая из них включает п. 1 и/или п. 2) для получения выборки последней ступени или конечных территориальных единиц.

4) Составление списка строений или домашних хозяйств в каждой отобранной конечной территориальной единице.

5) Отбор выборки таких единиц из списка.

6) Документальное оформление каждого из перечисленных выше шагов, как можно более подробное, с тем чтобы структура и процедуры выборки были вполне ясны и на основе обследования могли быть исчислены требуемые оценки и их дисперсия для всей совокупности. При продолжающихся обследованиях важно указать также пересечение единиц в выборках для различных раундов.

7) Подготовка списка выборочных домашних хозяйств для каждой конечной территориальной единицы, а также информации, необходимой для распределения работ и контроля за ходом опроса.

Когда имеются хорошие списки территориальных единиц, дополнительная работа в связи с шагами 1 — 3 может быть незначительной. Кроме того, если она необходима, эту работу можно провести задолго до начала основного сбора данных для обследования. Дело в том, что территориальные единицы относительно стабильны.

В отличие от этого, шаг 4, т.е. составление списка жилых единиц или домашних хозяйств, может представлять собой крупную работу до начала каждого раунда обследования. Как указано в парагр. 3 гл. 11, конечными единицами отбора в большинстве обследований рабочей силы являются домашние хозяйства или жилые единицы (адреса), а не компактные территориальные единицы или индивиды. Следует уделить пристальное внимание составлению списка конечных единиц, с тем чтобы обеспечить должный охват, а каждая отобранная единица могла бы быть идентифицирована на последующей стадии опроса. Составление списков и отбор выборки целесообразно организовать отдельно от основного опроса. Если составление списков и отбор единиц поручают тем же лицам,

которые занимаются опросом, возможны серьезные перекосы, особенно если все эти операции совмещены. Но важно также, чтобы разрыв во времени между составлением списков и опросом был минимальным. Стабильность списков зависит от обстоятельств и особенностей учтенных единиц; социальные единицы, подобно домашним хозяйствам, как правило, менее стабильны, нежели, например, здания и жилые единицы. Каковы бы ни были используемые единицы, существует вероятность, что некоторая работа по составлению списков потребуется до опроса в любом раунде обследования. Время и масштабы составления списков по отношению к основному опросу зависят от нескольких факторов. Во-первых, от доли отбора на последней ступени (т.е. отбора в конечных территориальных единицах), как это вытекает из плана выборки. Обычно требуется в три-шесть раз больше списочных единиц, нежели окончательное число опрашиваемых единиц. (Если соотношение намного больше названного, желательно разбить имеющиеся территориальные единицы на гнезда более подходящего размера.) Во-вторых, выборки для различных раундов обследования рабочей силы могут в той или иной степени пересекаться, что опять-таки зависит от плана; поэтому для любого раунда может потребоваться составление списка лишь для подмножества выборочных территориальных единиц. Однако, если одни и те же единицы используются в течение продолжительного времени, желательно их периодически обновлять с учетом стабильности списков. Условия меняются, но в качестве приблизительного ориентира списки домашних хозяйств могут сохраняться до 6—12 месяцев, а списки строений и жилых единиц — максимум 12—24 месяца. Стабильность списков влияет также на график их составления с учетом проведения основного опроса.

Наконец, трудоемкость составления списков зависит также от характера информации, собираемой в ходе этой операции. Список строений и жилых единиц, не требующих контактов с жителями, можно составить сравнительно быстро (скажем, из расчета 15—25 единиц в день на человека). Составление списков домашних хозяйств, как правило, предполагающее такие контакты, проходит медленнее и также в определенной степени зависит от собираемой информации. Значительно больше времени требуется при составлении списков индивидов и их характеристик. (Эти вопросы подробно обсуждены в World Fertility Survey, 1975a).

Во многих случаях удобно и желательно использовать различные группы персонала для составления списков и последующих опросов. Это может оказаться неизбежным, если обе операции по времени

пересекаются, как это происходит при непрерывном обследовании рабочей силы.

Отбор и документальное оформление выборки

Отбор единиц следует производить как можно более централизованным образом.

Он может оказаться довольно трудоемкой конторской работой, и ее компьютеризация способна дать значительные выгоды, особенно при крупномасштабных непрерывных обследованиях.

Следует подчеркнуть важность документального оформления отбора. Это тем важнее применительно к крупным обследованиям рабочей силы, которые могут служить основой для других обследований, и тогда, когда выборки следует увязывать и контролировать во времени. Желательно иметь три вида документов о плане выборки и его реализации.

1) Общее описание плана, отбора и результата.

2) Список конечных единиц отбора с характеристикой для каждой из них — идентификационного номера, структуры, единицы вышестоящей ступени, вероятность отбора, число записанных, отобранных и опрошенных единиц в территориальной единице и т.д. Такая информация может оказаться весьма полезной для контроля сбора и обработки данных. Она способствует также обеспечению полноты отбора. Ошибки отбора можно выявить только на основе полного описания структуры выборки.

3) Для каждой территориальной единицы, попавшей в отбор, — список конечных отобранных единиц (домашних хозяйств, жилых единиц) с необходимой для их идентификации информацией. Такие списки необходимы для распределения работ и контроля за интервьюерами.

5. СБОР ДАННЫХ

Ограничения по времени

При обследованиях рабочей силы операции по сбору данных особенно подвержены жестким ограничениям по времени. Дело в том, что важно обеспечить не только географическую представительность выборки, но и представительность во временном аспекте, чтобы можно было уловить сезонные и иные колебания, влияющие на экономическую активность (см. парагр. 5 гл. 11). Кроме того, при непрерывном обследовании мало возможностей отложить или продлить этап сбора данных любого раунда. Эти жесткие ограничения

по времени ведут к нескольким последствиям. Во-первых, крайне важно получить надежные оценки скорости работы интервьюеров. Эти показатели могут отличаться с учетом особенностей населения, и такие различия важно выяснить. Во-вторых, необходимы строгий контроль и надежная система связи, с тем чтобы обеспечить плановую производительность в работе интервьюеров. Где эта производительность не достигнута, необходимы срочные меры. В-третьих, система должна быть гибкой, чтобы можно было перераспределять работников и ресурсы, когда это необходимо. Важно также иметь в системе резервные мощности.

Организация, наблюдение и контроль

Общие принципы организации и контроля сбора данных на местах рассмотрены в работе ООН (1984, парагр. 5.34 — 5.86). Здесь мы выделим лишь основные моменты. Для контроля за движением документов и информации важно иметь какую-либо центральную административную единицу. В крупных странах этот орган можно дополнить региональными единицами. Необходимо хранить документацию с результатами опроса каждой единицы выборки. Надо обеспечить четкое и быстрое распространение материалов. Необходимо установить задания интервьюерам с учетом разнообразия условий; средние показатели производительности могут служить лишь в качестве общих ориентиров. Размер и способ оплаты труда интервьюеров может существенно повлиять на их поведение, мотивацию и качество работы.

Тщательный контроль за работой предполагает определенное соотношение числа контролеров и интервьюеров. Работа по контролю весьма разнообразна, но ее можно свести к нескольким основным моментам. 1) Заполненные интервьюерами анкеты следует проверить как можно раньше, особенно на ранних стадиях опроса. 2) Следует проверить часть анкет, чтобы удостовериться, что интервьюер работал только с единицами, включенными в выборку. 3) Где это возможно, контролер должен присутствовать на нескольких опросах, чтобы убедиться в качестве работы и в условиях проведения опроса. 4) Небольшую часть опросов можно повторить, чтобы проверить достоверность ответов и оценить работу отдельных интервьюеров. 5) Интервьюеров можно попросить записать часть работы на пленку. Это дает неоценимую информацию для оценки того, как проводят опрос. Запись может также служить хорошим подспорьем при подготовке и переподготовке интервьюеров (Woril Fertility Survey, 1975, section 6.7).

Распределение интервьюеров

Организация сбора данных — это еще один важный аспект, способный повлиять на качество работ и затраты. Существуют две взаимосвязанные проблемы: интервьюеры могут работать по одному или бригадами; их можно набирать из числа местных жителей для работы в определенном районе или же они могут быть мобильными, перемещаясь из одного района в другой. Разумеется, существуют различные степени мобильности. Возможны две крайние формы организации труда: а) использование весьма мобильных бригад, когда интервьюеры работают сообща и перемещаются вместе со своим контролером-бригадиром из одного района в другой по мере сбора данных; б) использование постоянных счетчиков, часто из числа местных жителей, каждый из которых работает в одиночку в определенном районе в течение длительного периода. В этом случае контролер может проживать в другом месте и периодически посещать каждого интервьюера. Между этими крайними случаями возможны промежуточные варианты.

Использование мобильных бригад, как правило, обеспечивает лучший контроль за ходом работы на местах. Оно может также способствовать получению более удачного плана выборки, так как мобильная бригада определенной численности может охватить большее количество выборочных районов: при обследованиях рабочей силы обычно желательна достаточно разбросанная и разнообразная выборка. Кроме того, типичное обследование не предполагает частых повторных опросов одних домашних хозяйств, как это делают при многих обследованиях потребления продовольствия и семейных бюджетов. Оба эти фактора уменьшают объем работы в одном районе и делают предпочтительной мобильность интервьюеров. Мобильная бригада позволяет также более гибко планировать ротацию выборки и другие аспекты плана, так как эти факторы делают нежелательным закрепление интервьюеров на местах. В пользу бригадного подхода говорит и фактор времени, о котором упоминалось выше, ибо такой подход позволяет легче маневрировать персоналом и замещать выбывших лиц. Основные трудности, связанные с мобильной бригадой, включают: большую потребность в транспортных средствах для каждой бригады; высокие транспортные расходы и затраты на жилье в каждом районе; возможна и более низкая производительность в связи с потерей времени на разъезды, которая может быть значительной, если члены бригады вынуждены ждать, пока последний не закончит свою работу. Серьезные затруднения при бригадном методе могут возникнуть также

в случае, когда имеются существенные различия культурного и языкового плана.

Преимущества следует сопоставить с недостатками. Иногда удачным компромиссом может служить использование небольших бригад (скажем, по два-три интервьюера в каждой из них) с ограниченной мобильностью. По сути дела, в одном обследовании можно прибегать к различным методам организации труда в разных районах.

6. ПОДГОТОВКА И ОБРАБОТКА ДАННЫХ

Операции обработки данных

Обработку данных часто называют "узким местом" обследования. Дело в том, что многие обследования пострадали от серьезного превышения сметы, крупных задержек и даже полной неудачи на этапе обработки данных. Часто это обусловлено нехваткой специалистов в области обработки данных, оборудования и программного обеспечения, а также неудовлетворительным руководством и контролем.

Следует различать два этапа этой работы, а именно, подготовку данных и собственно их обработку, хотя, как отмечалось ранее, различные этапы процесса все более интегрируются. Подготовка данных включает их ручную сверку и редактирование на местах и в конторе, присвоение численных кодов, документальное оформление и описание собранных данных, проверку на полноту учтенной выборки и, наконец, ввод данных в компьютер. Многие операции подготовки данных включают трудоемкие и повторяющиеся конторские работы, требующие большого числа оперативных работников. Проверка и другие меры по контролю качества их работы также должны быть обеспечены.

Последующий этап обработки данных связан с компьютерным редактированием и исправлением данных, внесением недостающих величин, изменением формы представления данных, получением новых показателей, табулированием и анализом и, наконец, хранением и распространением данных. Этот этап характеризуется большей зависимостью от наличия сложной аппаратуры и сравнительно небольшого числа системных аналитиков, программистов и других специалистов.

Всестороннее обсуждение этапа обработки данных обследования содержится в работе ООН (1982 г.), хотя некоторые моменты этого исследования нуждаются в обновлении с учетом быстрого прогресса в этой области. В следующем параграфе мы рассмотрим лишь два главных аспекта этой работы: кодирование; ввод данных и машинное редактирование (которые все чаще интегрируют).

Для начала полезно остановиться на некоторых специфических особенностях и требованиях к обработке данных обследований рабочей силы. 1) Если данные собирают в течение длительного периода или непрерывно, особенно важно обеспечить, чтобы не было отставания в обработке и чтобы она происходила теми же темпами, что и сбор данных. 2) Многие обследования рабочей силы рассчитаны на получение текущих показателей в строго установленные сроки: задержка в обработке и составлении отчета может серьезно отразиться на ценности информации. 3) В непрерывных обследованиях необходимо также обеспечить увязку данных отдельных раундов. Часто требуется увязка и между отдельными видами единиц, например, домашними хозяйствами и индивидами. 4) Табулирование и анализ нередко усложняются из-за необходимости взвешивать данные обследования по внешним контрольным итогам (см. гл. 11). 5) В отличие от данных более глубоких исследований, например обследований семейных бюджетов, данные обследований рабочей силы обычно проще по своей структуре и короче в расчете на выборочную единицу (домашнее хозяйство или индивид), но объем выборки, как правило, значительно больше. 6) В то же время многие обследования рабочей силы имеют то преимущество, что большинство вопросов можно заранее закодировать; занятие и отрасль — это часто единственные вопросы, требующие специального кодирования.

Кодирование

Кодирование — это операция, в ходе которой записям в анкете присваивают численные значения. Эта работа предполагает: а) разработку основы кодирования или множества взаимоисключающих категорий, охватывающих все приемлемые варианты ответов на вопросы; б) присваивание кода каждому варианту ответа. Сложность этой работы зависит от характера основы кодирования, числа подлежащих кодированию вариантов и взаимосвязи между этими двумя аспектами. В ряде случаев код состоит из нескольких легко различных элементов, в других случаях он может включать большое число трудноразличимых элементов. Точно так же некоторые ответы уже могут быть в числовой форме, которую лишь надо преобразовать в код, но в других случаях операция может оказаться значительно более трудной, сопряженной с ошибками и большими затратами времени, в частности, когда требуется кодировать словесные описания, не всегда ясные и полные, а сам код состоит из многих элементов. Очевидно, что при крупномасштабном обследовании желательно как можно больше ответов сделать закрытыми и

закодировать заранее. (Проблема структурирования вариантов ответов и сравнительных преимуществ открытых и закрытых вопросов рассмотрена в парагр. 4 гл. 12.)

В типичном обследовании рабочей силы наиболее важными и сложными для кодирования являются аспекты занятости, а именно, отрасль, занятие и в меньшей степени — статус занятого. Как правило, они предполагают многоаспектное кодирование на основе словесных описаний, сделанных в ходе опроса. Кроме того, в анкете могут содержаться полуоткрытые вопросы, в которых большинство вариантов ответов заранее закодированы, но допускаются другие ответы, которые либо неизвестны, либо редки, что не дает возможности составить исчерпывающий список и заранее его закодировать. В качестве примеров можно назвать вопросы о причинах работы меньше нормального времени, временного отсутствия на работе, отказа от поиска работы и т.д. Большинство вопросов, как правило, кодируют заранее.

Заранее закодированные вопросы

В заранее закодированных вопросах коды ответам следует присваивать на основе единых последовательных правил, например, “да” надо кодировать как “1”, “нет” — как “2”, другие ответы — как “8”, отсутствие ответа — “9” и т.д. Все ответы на один вопрос должны иметь код одинаковой значности. Обычно предпочтительнее кодировать каждый вопрос отдельно и не исключать информацию на этапе кодирования (United Nations, 1982, p. 12 — 13).

Важный вопрос заключается в том, следует ли заранее закодированные ответы переносить в другое место (на отдельную страницу анкеты или особый лист) для облегчения ввода данных. Анализ анкет обследований рабочей силы и других обследований показывает, что многие разработчики предпочитают переписывать ответы до ввода данных. Это бесспорно ускоряет ввод данных и может уменьшить число ошибок на этом этапе. Это может быть важным соображением, если ввод данных является более узким местом (ибо он требует специального оборудования), по сравнению с переписыванием (которое является чисто технической конторской работой). Можно, однако, утверждать, что переписывание излишне, ибо требует времени и может стать причиной новых ошибок в данных (United Nations, 1982, p. 12). Довод против переписывания получает новое звучание по мере развития интерактивных процедур ввода данных со встроенными контрольными механизмами (см. ниже).

Занятия и отрасли слишком многообразны и сложны, чтобы их можно было заранее закодировать. Коды включают много аспектов, которые не всегда легко распознать на основе словесных описаний, зарегистрированных в ходе опроса. Кодирование занятий и отраслей — это сложная операция, чреватая ошибками. Для этого требуются основательная подготовка и контроль качества. (Удачные эмпирические иллюстрации содержатся в работах: Jabine and Tepping, 1973; Lyberg, 1982).

На стадии проектирования можно принять некоторые меры с целью упрощения работы. Иногда полный открытый вопрос можно разделить на ряд вопросов, каждый из которых легче закодировать или даже сделать это заранее, хотя бы отчасти. Разделив полный вопрос на ряд вопросов, можно также повысить качество полученной информации. Так, занятие или отрасль можно сначала учесть по основной группе (возможно, в заранее закодированном виде), а затем — получив словесное описание подгруппы.

В ряде случаев кодирование отрасли можно облегчить, разделив работу на два этапа: а) на основе списка предприятий по отраслям закодировать как можно больше случаев по полученным наименованиям и адресам; б) оставшиеся случаи закодировать обычным путем на основе описания отрасли. Такой подход оправдан лишь тогда, когда занятые в достаточной мере сконцентрированы, что позволяет охватить большую долю случаев из выборки на этапе (а). При значительном пересечении выборок для отдельных раундов также удастся со временем составить списки предприятий по отраслям для указанной цели.

Желательно, чтобы в обследованиях рабочей силы профессия и отрасль кодировались по такой же схеме, как и в других крупных работах, например, общенациональных переписях населения. Разумеется, степень детализации в различных источниках может быть разной в зависимости от объема выборки.

Работу по подготовке данных можно организовать по-разному. Одна крайность — это полностью “плавающая” (свободная) система, при которой конторские служащие чередуют различные операции — сверку, переписывание, кодирование, контроль. Такая система предполагает всестороннюю подготовку персонала. При такой его подготовке свободная система имеет то преимущество, что работа становится менее повторяющейся и рутинной и может быть в целом более высокого качества. На другом полюсе существует “фиксированная” система, означающая, что каждый работник обу-

чен определенной операции и закреплен за ней. Преимущество этой системы в том, что каждый работник выполняет узкую операцию и вырабатывает навыки более эффективной работы. Это может оказаться более существенным применительно к кодированию сложных пунктов, например, профессии и отрасли. Поэтому при обследованиях рабочей силы для кодирования этих пунктов обычно используют специальных работников, тогда как для других операций подготовки данных прибегают к свободной системе.

Контроль работ и качества

В целях контроля операций и качества работы кодировщиков важна проверка (верификация). Под этим подразумевают повторение операции и сопоставление результатов двух операций для выявления частоты и источников ошибок. Обычно такую проверку достаточно осуществлять на выборочной основе, причем доля выборки зависит от доли выявленных ошибок. Но эта доля должна быть высокой (в ряде случаев даже 100%-ной) на начальных этапах работы, применительно к более сложным пунктам и операторам, которые, как было установлено, допускают больше ошибок. (Оценка и контроль кодирования и других ошибок при обработке данных более подробно обсуждаются в гл. 14.)

Верификация может быть зависимой и независимой. Под зависимой верификацией понимают процедуру, когда проверяющий имеет доступ к результатам предыдущей операции. Он просматривает проставленный код и решает, правилен ли он, а затем исправляет, если нужно, по своему усмотрению. Зависимую верификацию можно осуществить быстрее и дешевле, но она может и не выявить все ошибки, требующие исправления. Проверяющий должен также иметь более высокую квалификацию, чем первоначальный кодировщик. При независимой верификации операцию выполняют менее двух работников, а в случае расхождений выявляют и исправляют ошибки. Это можно осуществить путем дополнительного проведения операции с последующим принятием решения большинством голосов, или же правильный код определяет специалист. Независимая верификация позволяет скорее выявить ошибки и удобна в том смысле, что проверку могут осуществлять рядовые кодировщики. Но она требует значительного времени, а обеспечить независимость верификации от кодирования может оказаться нелегко. Можно также предусмотреть *оценку* в отличие от верификации. Оценка предполагает существование "истинных" кодов, с которыми можно сравнить первоначальные коды с целью выявить долю

ошибок. Она требует опытных кодировщиков и основана на предположении, что они могут определить "истинные" коды.

В ряде случаев возможно автоматическое кодирование (с помощью компьютера). Однако этот вариант, как правило, недопустим в обследованиях рабочей силы в развивающихся странах.

Ввод и редактирование данных

Традиционный ввод данных — это сама по себе простая механическая операция, почти не требующая особых знаний об обследовании и его содержании. Однако неоднократно установлено, что это серьезный источник задержки в крупномасштабных обследованиях из-за недостаточных мощностей (оборудования и/или операторов) для выполнения этой операции.

В результате прогресса в компьютерной технологии ситуация в области ввода данных быстро меняется. Во-первых, это прогресс в оптических устройствах для считывания значков или букв. Ввод данных на основе оптического сканирования позволяет исключить работу оператора в этой части, экономит время и сокращает число ошибок ввода. Однако использование этих методов требует дорогостоящего оборудования и его обслуживания, соответствующего построения и полиграфического исполнения анкет, а также аккуратного обращения с документами на местах и в конторе, а все это проблематично при проведении обследований в развивающихся странах (United Nations, 1982, p. 20).

Более существенные перемены связаны с развитием интерактивного ввода данных. Это означает, что в процессе ввода данных автоматически контролируются и редактируются их структуры и формат. Многие статистические организации развивающихся стран уже накопили значительный опыт применения интерактивного ввода данных, и этот опыт ширится по мере распространения персональных компьютеров и соответствующего программного обеспечения. Наиболее важный аспект этого процесса — это, пожалуй, возможность сократить время компьютерного редактирования. Позволяя раньше выявлять ошибки, новые методы уменьшают также необходимость редактирования вручную.

Организация ввода и редактирования данных

Во многих случаях ввод и редактирование данных следует организовать в виде сравнительно централизованной операции в штаб-квартире обследования, а в крупной стране, возможно, — в регио-

нальных или провинциальных офисах. (Однако при наличии достаточных средств в развивающихся странах недавно прибегали к более децентрализованной системе интерактивного ввода данных.) Важно установить должную систему получения и контроля анкет и других материалов. Обычно для этой цели удобно использовать конечные территориальные единицы выборки в качестве базовых пунктов: все анкеты из данной единицы поступают на обработку одновременно, проверяются на полноту в соответствии со списками выборочных домашних хозяйств (или жилых единиц) для данной территории, а затем хранятся вместе для кодирования, ввода и редактирования. При обследовании рабочей силы каждое домашнее хозяйство будет иметь одну общую анкету и, возможно, несколько индивидуальных анкет для соответствующих индивидов. Важно проверить, получены ли все анкеты от каждого выборочного домашнего хозяйства. Рекомендуется вести два типа реестров: подробный реестр таких элементов, как число домашних хозяйств в каждой выборочной жилой единице, число полученных анкет домашних хозяйств, число лиц, подлежащих опросу в каждом домашнем хозяйстве; число полученных индивидуальных анкет и т.д.; а также территориальный реестр, обобщающий перечисленную информацию применительно к каждой конечной территориальной единице. Затем в процессе редактирования полезно также вести счет ошибкам по видам, а также исправлений и добавлений отсутствующих данных.

Ручное и машинное редактирование

В ходе обследования данные обычно редактируют несколько раз: проверка интервьюером и бригадиром на месте; редактирование вручную в конторе до кодирования; машинное редактирование после ввода данных. На каждом этапе, по сути дела, проверка сводится к одному и тому же, однако степень детализации может быть разной. Проверки на местах неизбежны, ибо обычно только они позволяют повторно выйти на респондента для уточнения полученных данных. Важная проблема состоит в выборе между редактированием вручную (в конторе) и машинным редактированием. В связи с ограниченными ресурсами для машинного редактирования и более широкими возможностями привлечения конторского персонала для редактирования вручную во многих обследованиях обычно подчеркивают роль последнего. При этом исходят из того, что машинное редактирование не является основным в деле выявления и исправления ошибок, а его задача — обна-

ружить ошибки, оставшиеся незамеченными на ручной стадии, и ошибки, появившиеся при кодировании и вводе данных. В результате улучшения компьютерной технологии и особенно развития интерактивного ввода данных ситуация в целом меняется. Но хотя некоторые специалисты утверждали, что от этапа контрольной проверки можно вообще отказаться, многие согласны с тем, что роль этого этапа, возможно, и менее значительная, сохраняется, по меньшей мере в части должной подготовки анкет для кодирования и устранения грубых ошибок до ввода данных и машинного редактирования. Указанные соображения, очевидно, еще более весомы применительно к обследованиям, в которых анкеты сравнительно сложны, а объем выборки — небольшой или умеренный.

Компьютерное редактирование

Традиционное компьютерное редактирование предполагает ряд шагов: *редактирование формы*, при котором проверяется идентификация единицы, размещение данных и наличие нечисловых символов и т.д.; *структурное редактирование* — проверка данных на полноту, т.е. наличие всех выборочных единиц и всех записей для каждой из них; *контроль значности* (диапазона), т.е. выяснение того, находятся ли значения величин в допустимых пределах; *проверка переходов*, т.е. контроль последовательности постановки вопросов и отсутствия информации, не относящейся к данному лицу; наконец, проводят ряд *проверок непротиворечивости* данных, т.е. их соответствия друг другу. Могут быть также выявлены факты отсутствия данных, которые восполняют на этом этапе. Обычно эти шаги выполняют последовательно и, как правило, они требуют обращения к оригинальным документам обследования и исправлений вручную, а весь процесс проходит ряд циклов. Трудности, возникающие при машинном редактировании, стали главной причиной задержек во многих обследованиях. Нередко это связано с тем, что выделяют недостаточно работников для обращения к анкетам, выявления и исправления ошибок, показанных на распечатках, и обновления массивов данных. Часто эту операцию приходится выполнять неоднократно для каждого цикла процесса редактирования (см. Otto and Rattenbery, 1987). Интерактивный ввод данных призван интегрировать эти различные шаги и существенно облегчить и ускорить обработку данных.

Охарактеризованные выше проверки требуются конкретизировать, обеспечить программно и применить в деталях на всем масси-

ве. Хотя небольшое число ошибок или отсутствующих данных могут и не повлиять существенно на *статистическую* значимость агрегированных результатов обследования, тем не менее желательно устранить все несоответствия, и на это существует несколько причин. Наличие ошибок обычно существенно затрудняет обработку данных, программирование и анализ. Это может вызвать особые проблемы, когда необходима увязка данных с данными из других источников, требуется предпринять обширный анализ, и особенно, когда работа рассчитана на различных аналитиков и пользователей. Приведенные соображения приобретают особую важность по мере того, как ширится распространение данных в различных формах, а не только в виде агрегированных таблиц. Разумеется, редактирование и исправление ошибок могут быть весьма трудоемким делом, поэтому требуется сопоставить затраты и практическую пользу от этой работы.

Спецификация редакционных проверок. Некоторые примеры

В последней части параграфа приведены некоторые примеры и дополнительные описания отдельных видов контроля данных, обычно применяемого при обследованиях рабочей силы. Разумеется, конкретные формы контроля будут зависеть от состава анкеты. Тем не менее эти примеры могут оказаться полезными при разработке форм контроля для различных ситуаций.

1) *Редактирование формы.* Данные обследований обычно расположены в виде иерархии составных частей и по единицам выборки, записям внутри каждой единицы и т.д., причем каждому элементу присваивают свой идентификационный номер. Редактирование формы предполагает проверку этих элементов и наличие правильных идентификационных номеров. В то же время можно осуществить общий контроль на наличие нечисловых и других неверных знаков. Обычно сортировку производят после контроля и корректировки идентификационных признаков.

2) *Редактирование структуры.* Главная цель — проверить, имеются ли данные по каждой выборочной единице, а внутри каждой единицы — насколько полна и выдержана структура. К примеру, в обследованиях рабочей силы контроль призван установить наличие:

а) некоторой информации по каждому выборочному домашнему хозяйству, что должно по меньшей мере показать, насколько успешно был проведен опрос;

б) по каждому опрошенному домашнему хозяйству — всех элементов данных (записей и т.п.) в целом по домашнему хозяйству и его членам (а также выяснить, следует ли в деталях определять экономическую активность каждого лица);

в) некоторой информации по лицам, подлежащим опросу, по меньшей мере указывающей на то, что это лицо было успешно опрошено;

г) по каждому опрошенному лицу — элементов данных, включенных в опрос.

3) *Контроль значности.* Здесь следует отметить несколько практических моментов. По всем вопросам значность ответов уже должна была быть точно указана в анкете и/или руководстве для кодирования. При спецификации этого контроля полезно прежде всего выдать предельные распределения данных, с тем чтобы можно было выявить наличие и частоту выхода за пределы рабочего диапазона в каждом поле данных.

Вопросы, не относящиеся ко всем респондентам, должны содержать код “не применимо” (НП); это не распространяется на вопросы, имеющие отношение ко всем респондентам. В ходе этого контроля можно выявить отсутствующие значения (“не указано” — НУ), проанализировать их частоту и распределение. Применительно к некоторым вопросам (например, доход, продолжительность занятости и безработицы, отработанные часы и т.п.) целесообразно определить более узкий, нежели теоретически возможный диапазон, с тем чтобы выявить необычные (а следовательно маловероятные, хотя и не обязательно невозможные) значения, которые, по всей вероятности, могут оказаться ошибочными.

4) *Контроль переходов.* Его цель — убедиться, что опрос протекал правильно: на каждый уместный вопрос получен ответ (даже если это НУ), а все НП должны быть пропущены или выявлены еще каким-либо образом. Контроль переходов можно существенно ускорить с помощью логической схемы анкеты, которая, как отмечено в парагр. 2 гл. 12, показывает взаимосвязь между вопросами и подгруппами населения, к которым относится каждый вопрос. Контроль переходов — это попросту логические условия, определяющие, когда уместен тот или иной вопрос. Ниже приведен пример на основе логической схемы 11 из парагр. 3 гл. 12. Обратите внимание, что решение относительно перехода следует принимать по каждому вопросу, если ответ НУ; все действительные ответы (включая НУ) необходимо отразить во всех случаях.

ПРИМЕР КОНТРОЛЯ ПЕРЕХОДА НА ОСНОВЕ ЛОГИЧЕСКОЙ СХЕМЫ II

Вопрос	Применимо, если и только если:
B1	(всегда)
B2	B1 — нет, НУ
B3	B2 — нет, НУ
B4	B3 — нет, НУ
B5	B4 — домашняя работа, прочее, НУ
B6-8	B4 — пенсионер или B5 — да
B9	4 — пенсионер или B5 — нет, НУ
B10	B9 — да
B11	B1 — нет
B12, 16, 18-22	B1 — да или B2 — да
B13	B12 — да
B14, 15	B12 — нет, НУ
B17	B16 — неполное время, НУ
B23	B22 — работодатель, наемный работник, занят на собственном предприятии, НУ
B24-27	B3 — да
B28	B27 — да

5) *Смежный контроль.* Проверку последовательности и полноты опроса можно осуществить также иначе. Иногда ответ на вопрос можно зарегистрировать (или обобщить) на основе одного или нескольких предшествующих вопросов, проверив ответ на непротиворечивость. Еще одна распространенная форма контроля — это введение в анкету таблиц (т.е. регистрация подробностей о различных прежних или нынешних работах), причем число записей в таблице должно соответствовать ответу на какой-либо предыдущий вопрос. (Названные выше виды контроля обычно именуют соответственно “фильтрующим” и “табличным”).

6) *Проверка на непротиворечивость.* Речь идет о совместимости ответов на различные, но связанные между собой вопросы. Вид возможной и необходимой проверки зависит от содержания и построения анкеты. Большой контроль возможен, когда в полученной информации содержится определенная избыточность либо различные вопросы тесно связаны или пересекаются (как это, например, имеет место в большинстве обследований рабочей силы). Многие более простые анкеты построены с использованием указаний о переходах, главным образом с целью избежать несовместимых ответов. Тем не менее некоторый контроль внутренней непротиворечивости возможен и полезен, как это видно из приводимого ниже примера.

ПРИМЕР ПРОВЕРКИ НА НЕПРОТИВОРЕЧИВОСТЬ НА ОСНОВЕ ЛОГИЧЕСКОЙ СХЕМЫ 11

Вопрос	Желательность проверки в отношении:
B4, 6, 13, 28	возраст индивида
B7, 19, 26, 28	образовательный уровень индивида
B23	профессия индивида (B19)

Обратите внимание, что эти проверки дополняют контроль на значность (диапазон) и другие проверки применительно ко всем вопросам в отдельности.

7) *Исправление, добавление и отсутствие информации.* Как только выявлена ошибка, ее можно исправить одним из следующих способов:

а) если вопрос в данном конкретном случае неприменим, данные исключают;

б) ошибочные (или отсутствующие) данные исправить нельзя, и их заменяют на НУ;

в) ответ исправляют (или добавляют) с учетом другой информации, содержащейся в анкете;

г) неправильные или отсутствующие данные вносят на основе информации, внешней по отношению к анкете, например, из предыдущего раунда опроса той же единицы или "схожих" единиц выборки.

Разумеется, ответ можно оставить без исправлений, но, если доля ошибочных ответов невелика, обычно предпочтительнее прибегнуть к варианту (б), если ничего лучшего сделать нельзя. Этот вариант удобен также для прекращения процедуры контроля, когда польза от него уже незначительна при сравнении с затратами на дальнейшее редактирование.

Заметьте, что часто нелегко провести четкое различие между правильными (непригодными) и отсутствующими данными или между исправлением (нередко произвольным) и добавлением данных. Добавление обычно означает исправление на основе информации, внешней по отношению к данной анкете. Вообще говоря, предпочтительнее вносить изменения и коррективы в данные обследований вручную, на основе тщательного рассмотрения каждого случая и обращения к оригинальным анкетам. Добавление данных, пожалуй, больше поддается автоматизации, особенно применительно к крупным и не очень сложным по содержанию данным. В любом случае процесс исправления и добавления требует тщательного контроля, а все изменения следует зарегистрировать.

Составление отчета

При анализе и составлении отчета о результатах опроса обычно удобно разделить работу на два основных этапа. Первый может состоять в выпуске главных комбинационных таблиц и комментировании основных выводов. На этом же этапе можно выпустить ряд ответов, не требующих пространного и сложного статистического анализа. Каждый из них должен быть ориентирован на определенную категорию пользователей. Эти документы могут включать: предварительный и краткий отчет о главных выводах, даже еще до того, как получены вполне “чистые” и детальные файлы и таблицы; обобщенные характеристики ряда конкретных вопросов и проблем с ориентацией на определенных пользователей; технический отчет, описывающий методологию и процедуры обследования.

Второй этап может включать более развернутую программу анализа и исследований с использованием более специальных и сложных методов. Обычно он предполагает сотрудничество со специалистами и организациями вне статистического бюро. Выпуск, хранение и рассылка должным образом оформленных файлов микроуровня существенно облегчают эту работу. При анализе следует стремиться увязать результаты обследования с аналогичной информацией из других источников, а особенно — с предыдущими раундами обследования рабочей силы. Можно также предпринять более детальный анализ качества данных, источников ошибок, полученных методологических выводов и т.д. Особое значение имеет расчет ошибок выборки.

Оба этапа анализа и отчетности различаются по срокам и целям. Первый этап имеет целью обнародовать информацию об обследовании и его основных результатах как можно скорее. Второй этап ориентирован на более полное использование результатов при менее жестких ограничениях по срокам и при более широком участии специалистов. Он способствует также расширению аналитических возможностей статистического органа.

Табулирование

Первый этап отчета может базироваться на многоаспектных комбинационных таблицах. При табулировании следует использовать данные обследования как можно полнее, насколько это допускает наглядность. В качестве иллюстрации в конце параграфа приведен минимальный набор таблиц на базе обследования рабочей силы.

Этот перечень можно, разумеется, видоизменить и дополнить в конкретных обследованиях в зависимости от полученной информации. Ниже приведен также список основных показателей, отражаемых в таблицах. До выпуска таблиц часто целесообразно сочетать или перестроить “сырые” данные в виде *перекодированных показателей*, которые прямо отражаются в таблицах. В качестве примера можно сослаться на рис. 6 и 7 парагр. 4 гл. 6, где показано, как можно вывести показатель текущего состояния рабочей силы на базе серии вопросов об активности в данный момент, причинах отсутствия на работе, шагах по ее поиску или причинах отказа от поиска, готовности приступить к работе в данный момент, прошлом опыте работы и т.д. Если вывести такой показатель раз и навсегда для каждого индивида в выборке, то это существенно ускорит последующие табулирование и анализ.

Хранение данных и рассылка

С целью лучшего использования данных важно, чтобы они были доступны на микроуровне в виде должным образом оформленных компьютерных файлов. По мере прогресса компьютерной технологии такая форма рассылки данных приходит на смену традиционным комбинационным таблицам как “конечному” продукту обследования. Национальные органы статистики должны стремиться создать компьютерные банки данных или архивы, ориентированные на пользователей. Основные предпосылки при этом таковы:

- файлы данных отредактированы и оформлены должным образом, а также структурированы так, чтобы ускорить анализ;

- файлы данных на микроуровне снабжены исчерпывающим описанием данных (имеются списки кодов, распределения значений ответов, итоги, суммы и т.д., желательно в машиночитаемой форме);

- имеются базы данных, увязывающие информацию из различных раундов и смежных источников (по меньшей мере такая информация накоплена);

- разработана система управления данными, их контроля и рассылки;

- пользователи данных могут рассчитывать на техническую и компьютерную поддержку, когда она требуется.

Отличным примером служит “Динамическая база данных” Международного института статистики, который хранит и рассылает файлы данных большого числа демографических обследований. Хорошее техническое описание конкретных применений содержится в работе (Rowe and Croft, 1987).

Примерный перечень основных показателей и таблиц обследования рабочей силы

Представленный ниже примерный перечень основных показателей и таблиц разработан с учетом: а) рекомендаций, содержащихся в международных стандартах по статистике экономически активного населения (ICLS, 1982); б) обеспечения совместимости с перечнем таблиц об экономических характеристиках, рекомендуемых для переписей населения (United Nations, 1986). Примерный перечень содержит 47 таблиц, характеризующих различные группы населения и домашних хозяйств в разрезе 25 основных показателей.

Восемь показателей представляют собой демографические и образовательные характеристики индивидов или домашних хозяйств. Это — пол, возраст, семейное положение, посещение школы, уровень образования, отношение к главе домашнего хозяйства, размер домашнего хозяйства и количество детей.

Другие 17 показателей характеризуют занятость индивидов или домашних хозяйств. Это — статус активности в данный момент; обычный статус активности; количество недель или дней, отработанных на всех работах в истекшем году; основное (последнее) занятие; основная (последняя) отрасль; основной (последний) статус занятого; главная причина отсутствия активности лиц неактивных в настоящее время; функциональные категории лиц обычно неактивных; характер связи с рынком труда экономически неактивных лиц; количество дней или часов, отработанных на всех работах за прошлую неделю; причины отсутствия на работе; количество дней или часов, которые индивид готов дополнительно работать в неделю; шаги, предпринимаемые для поиска работы; причины отказа от поиска работы; число экономически активных лиц; число занятых и число безработных лиц в домашнем хозяйстве.

Для целей международных сопоставлений соответствующие категории для каждого показателя должны, где это возможно, отвечать международным стандартным классификаторам или быть доступны для преобразования с учетом классификаторов. К таким классификаторам относятся:

- а) Международный стандартный классификатор профессий — Международное бюро труда;
- б) Международный стандартный классификатор отраслей всех видов экономической деятельности — ООН;
- в) Международный классификатор статуса занятого (работодатель, наемный работник и т.д.) — определения даны ООН — за исключением классификации неоплачиваемых членов семьи, для

которых критерий минимального времени (по меньшей мере треть нормального рабочего времени) более неприменим;

г) Временные рекомендации по классификации возрастных групп — ООН.

В приведенном ниже примерном перечне таблиц показатель в скобках говорит о том, что возможна группировка по этому показателю, а его использование зависит от точности данных в соответствующих клетках с учетом уменьшенного объема выборки, на которых основаны данные клетки. Это общее соображение следует учитывать во всех комбинационных группировках, с тем чтобы избежать опубликования таблиц с клетками, данные в которых крайне неоднородны по выборке.

В ряде таблиц показатель или группа населения даны в скобках. Это указывает на то, что соответствующие таблицы можно составить применительно к обычному статусу активности в виде альтернативы или дополнения к статусу активности в данный момент: табл. 1, 4-6 и 43—47 можно составить применительно к статусу активности в данный момент, обычному статусу активности или тому и другому статусу. Табл. 2 составляют, если при обследовании учитывают как статус в данный момент, так и обычный статус. Наконец, табл. 15—42 о занятости, видимой неполной занятости и безработице предполагается составлять только применительно к статусу активности в данный момент.

Перечень основных таблиц о занятости, безработице и неполной занятости экономически активного населения

Население трудоспособного возраста (Население в возрасте ... лет и старше)

- 1) статус активности в данный момент (обычный статус активности), в разрезе половозрастных групп
- 2) статус активности в данный момент и обычный статус активности, в разрезе половозрастных групп
- 3) обычный статус активности и число недель или дней, отработанных на всех работах в прошлом году, в разрезе полов

Молодежь трудоспособного возраста

- 4) статус активности в данный момент (статус обычной активности) и посещение школы, в разрезе половозрастных групп

Население, активное в данный момент (обычно активное население) в разрезе:

- 5) семейного положения и половозрастных групп

- 6) основного занятия и половозрастных групп
 - 7) основных отраслей и половозрастных групп
 - 8) основного статуса занятого и половозрастных групп
 - 9) основной отрасли, основного занятия и полов
 - 10) основного занятия, образовательного уровня и половозрастных групп
 - 11) основной отрасли, образовательного уровня и половозрастных групп
 - 12) основного занятия, основного статуса занятого и полов
 - 13) основной отрасли, основного статуса занятого и полов
-

Женское население, активное в данный момент (обычно активное женское население)

- 14) в разрезе семейного положения и основного статуса
-

Лица, неактивные в данный момент (лица, обычно неактивные)

- 15) в зависимости от главной причины неактивности (функциональные категории) в разрезе половозрастных групп
 - 16) в зависимости от характера связи с рынком труда
-

Занятые, в разрезе:

- 17) основного занятия и половозрастных групп
 - 18) основной отрасли и половозрастных групп
 - 19) основного статуса занятости и половозрастных групп
 - 20) основной отрасли, основного занятия и полов
 - 21) основного занятия, основного статуса занятого и полов
 - 22) основной отрасли, основного статуса занятого и полов
-

Занятые и работающие, в разрезе:

- 23) основного занятия, числа дней или часов, отработанных на всех работах за последнюю неделю, и полов
 - 24) основной отрасли, числа дней или часов, отработанных на всех работах за последнюю неделю, и полов
 - 25) основного статуса занятого, числа дней или часов, отработанных на всех работах, и полов
-

Занятые, но отсутствующие на работе

- 26) в зависимости от причины отсутствия и в разрезе полов
-

Неполностью занятые (видимый аспект), в разрезе:

- 27) количества дней или часов готовности для дополнительной работы и половозрастных групп
- 28) основного занятия и половозрастных групп

29) основной отрасли и половозрастных групп

30) основного статуса занятого и половозрастных групп

Совместители, в разрезе:

31) половозрастных групп

32) основного занятия и полов

33) основной отрасли и полов

34) основного статуса занятого и полов

Безработные

35) в зависимости от отношения к главе домашнего хозяйства и в разрезе полов

Безработные, имеющие опыт работы, в разрезе:

36) последних занятий и полов

37) последних отраслей и полов

38) последнего статуса занятого и полов

Безработные без опыта прошлой работы, в разрезе:

39) половозрастных групп

40) полов и образовательного уровня

Безработные, активно ищущие работу

41) в зависимости от предпринятых шагов и в разрезе полов

Безработные, не ищущие активно работу

42) в зависимости от причины отказа от поиска и в разрезе полов

Глава домашнего хозяйства

43) в зависимости от статуса активности в данный момент (обычной активности) в разрезе половозрастных групп

Домашние хозяйства и их члены

44) в зависимости от размера домашнего хозяйства и числа экономически активных членов (с учетом активности в данный момент и/или обычной активности)

Домашние хозяйства с хотя бы одним безработным

45) по числу детей и занятых индивидов (с учетом активности в данный момент и/или обычной активности)

46) по числу безработных (с учетом активности в данный момент и/или обычной активности) в разрезе полов и основной отрасли работы главы семьи

Домашние хозяйства и члены домашних хозяйств, главы которых активны в данный момент (обычно активны)

47) в зависимости от основной отрасли, основного статуса занятого и пола главы семьи

Литература

- International Conference of Labour Statisticians (ICLS). 1982. "Resolution concerning statistics of economically active population, employment, unemployment and underemployment" (text also in French and Spanish), in *Bulletin of Labour Statistics* (Geneva, ILO), 1983—3, pp. xi-xvi.
- Jabine, T.; Tepping, B. 1973. "Controlling the quality of industry and occupation data", in *Bulletin of the International Statistical Institute*, 45 (3), pp. 360—389.
- Lyberg, L. 1982. "Coding of occupation and industry: Some experiences from Statistics Sweden", in *Bulletin of Labour Statistics* (Geneva, ILO), 1982—3.
- Otto, J.; Rattenbury, J. 1987. "WFS data processing strategy", in Cleland and Scott (eds.): *The World Fertility Survey: An assessment*. Oxford, Oxford University Press.
- Rowe, B.; Croft, T. 1987. "Construction of analysis files", in Cleland and Scott, op. cit.
- United Nations. 1982. *Survey data processing*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.
- 1984. *Handbook of household surveys*. Series F, No. 31, New York, UN Statistical Office.
- 1986. *Draft supplementary principles and recommendations for population and housing censuses*. Statistical papers, Series M, No. 67, New York, UN Statistical Office.
- World Fertility Survey. 1975. *Survey organisation manual*. WFS Basic Documentation, No. 2. The Hague, International Statistical Institute.
- 1975a. *Manual on sample design*. WFS Basic Documentation, No. 3. The Hague, International Statistical Institute.
- 1976. *Training manual*. WFS Basic Documentation, No. 4. The Hague, International Statistical Institute.

Глава 14. ТОЧНОСТЬ И ОЦЕНКА ДАННЫХ

«Подобно множеству других данных, информация об экономически активном населении, занятости, безработице, неполной занятости и по иным смежным вопросам подвержена ошибкам. Программы сбора данных следует тщательно планировать с целью минимизации возможных ошибок, однако некоторые из них неизбежны. Поэтому всесторонняя оценка результатов предполагает знание качества данных. Знание качества данных необходимо также для улучшения процедур сбора, обработки и оценки данных... Оценка качества должна, насколько это возможно, быть составной частью самой программы сбора данных.

При любом обнародовании статистической информации следует охарактеризовать качество данных, включая ошибки выборки и прочие ошибки...»

(Резолюция 1 13-й МКСТ, 1982, парагр. 37, 39 (ILO, 1983))

1. ПРИНЦИПЫ ВЫЯВЛЕНИЯ ОШИБОК

Все статистические данные, независимо от их источника и метода получения, потенциально подвержены разного рода ошибкам. Поэтому результаты переписей и обследований сопровождаются описанием их качества и недостатков.

Во-первых, знание качества данных необходимо для их *правильного использования и истолкования*. Это знание существенно для определения того, реальны ли и в какой степени заслуживают доверия их результаты и не отражают ли они лишь изменчивость и изъяны, присущие данным. Информация о характере и масштабах ошибок может также быть полезной для соответствующей корректировки данных и внесения поправок в их толкование.

Во-вторых, характеристика качества данных важна для оценки и улучшения плана и процедур обследования. Детальное изучение источников, размеров и влияния ошибок необходимо для выработки путей улучшения плана и процедур обследования и более эффективного распределения ресурсов между различными аспектами работ.

Постоянное слежение за качеством полученных данных и их улучшение особенно важны применительно к такому крупному начинанию, как общенациональное обследование рабочей силы. Результаты такого обследования часто образуют основу для принятия важных политических решений; они могут также стать исходным пунктом для сопоставления с информацией из многих дру-

гих источников. Как указано в предыдущих главах, результаты могут быть весьма чувствительны к используемым понятиям и определениям, к тому, как они вошли в работы, а также условиям сбора и обработки данных.

Виды ошибок

В настоящей главе охарактеризованы методы измерения различных аспектов ошибок в обследованиях. Поэтому полезно кратко описать виды и источники ошибок. Цель при этом состоит не в том, чтобы внушить отрицательное представление о результатах выборочного обследования как такового. Скорее она заключается в том, чтобы подчеркнуть важность принятия мер для контроля и уменьшения ошибок и должного их учета, если они неизбежно остаются при интерпретации и использовании результатов обследования.

Цель выборочного обследования — получить оценки и выводы, пригодные для всей совокупности, на основе наблюдения за ограниченным числом (выборкой) единиц совокупности. Можно провести различие между двумя группами ошибок, влияющими на процесс:

а) ошибки, возникающие в ходе экстраполяции результатов, полученных на основе единиц выборки, на всю изучаемую совокупность. Сюда входят ошибки охвата, отбора и реализации выборки, неполучение ответов, а также изменчивость выборки и смещение оценок. Эта группа ошибок связана с процессом планирования и реализации выборки, а также оценки результатов;

б) ошибки, обусловленные тем фактом, что то, что наблюдается и измеряется, не совпадает с тем, что предполагалось измерить на основе единиц, включенных в обследование. Это ошибки измерения, они связаны с содержательным аспектом обследования: определением его целей, преобразованием последних во взаимоувязанную и последовательную серию вопросов, их доведением до респондентов, способностью и желанием респондентов дать искомую информацию, а также регистрацией, кодированием, редактированием и обработкой ответов.

Приведенная классификация ошибок базируется на практических соображениях, и она в каком-то смысле более основательна, нежели обычное деление на ошибки выборки и прочие ошибки. Группа (а), затрагивающая распространение данных наблюдений на всю совокупность, включает изменчивость отбора, а также различные смещения, связанные с отбором и реализацией выборки, например, ошибки ответа и неполучения ответов. Все они имеют прямое отношение к плану выборки. Нередко несколько обследо-

ний и раундов обследования базируются на одной основе выборки, главной выборке, плане выборки и даже одних и тех же ее элементах. В этом случае ошибки, присущие процессу отбора единиц, как правило, становятся присущими для всех этих обследований, почти независимо от их содержания, поэтому возможны и общие для таких обследований методы оценки и контроля ошибок. Сказанное, разумеется, относится и к различным раундам продолжающегося обследования рабочей силы.

К группе (б) относятся ошибки в точности измерений на уровне индивидуальных единиц, учитываемых при обследовании. Они связаны с тем, насколько данные, сообщенные респондентом, зарегистрированные интервьюером, закодированные и проверенные конторским работником, исправленные, добавленные и табулированные компьютером, отличаются от действительной ситуации данного индивида. Эту группу ошибок можно изучить применительно к различным этапам обследования: сбору данных (получению ответов), кодированию, редактированию, добавлению отсутствующих показателей и обработке на компьютере.

Приведенная выше общая классификация должна быть дополнена перечислением *конкретных* источников и видов ошибок, как можно более полным, с тем чтобы облегчить контроль за ошибками и оценку их последствий. К примеру, среди ошибок реализации выборки полезно различать не только ошибки неполного охвата и неполучения ответов, но и выделить неполный охват домашних хозяйств; или различать неполучение ответов в результате отказа либо в связи с невозможностью выявить респондентов и войти с ними в контакт.

Вообще говоря, влияние любого источника ошибок на оценки обследования можно разделить на две компоненты: *ошибка изменчивости* и *смещение*. Некоторые условия проведения обследования "существенны" для его реализации; речь идет об общих социальных условиях, качестве имеющейся основы выборки, принятом методе сбора данных, построении анкеты, квалификации и подготовке персонала, условиях их труда, включая характер найма, уровень и метод оплаты; характеристике респондентов и т.д. Кроме того, на результаты обследования влияют также преходящие и случайные факторы, например, характер отобранных единиц выборки, интервьюеров и кодировщиков, условия проведения конкретного опроса и т.п. При данных "существенных" условиях возможны, следовательно, различные варианты проведения обследования, но они дали бы различные результаты в силу именно преходящих и случайных факторов (Hansen et al., 1961). *Ошибка изменчивости* и характеризует колебания оценок при гипотетическом повторении обследо-

ния. Средняя из возможных повторений — это *ожидаемая* величина оценки при данных существенных условиях проведения обследования. Разность между этим ожидаемым, или средним, значением и истинным значением показателя для всей совокупности и есть *смещение*.

Ошибки охвата и связанные с этим ошибки отбора и реализации выборки, неполучения ответов и ошибки систематического характера в ответах — это, в общем, смещения, обусловленные существенными условиями проведения работы. При повторении обследований в этих же условиях можно ожидать ошибок сопоставления с информацией, внешней по отношению к самому обследованию и полученной с помощью иных, более точных методов.

В табл. 20 перечислены основные виды ошибок обследования. Первая группа — “ошибки наблюдения” характеризуют разность между полученными в результате наблюдения и истинными значениями показателей для отдельных единиц обследования. Она подразделена на три подгруппы в зависимости от стадии, на которой возникают ошибки: 1) концептуальные ошибки на стадии планирования; 2) ошибки на стадии получения ответов при сборе данных; 3) ошибки при обработке данных, которые можно разделить на ошибки редактирования, кодирования, ввода данных и т.д. Особенно многообразны ошибки в ответах. Некоторые из них имеют характер систематических смещений, другие носят более случайный характер (например, “разногласия интервьюеров”). Следующая группа — “ошибки экстраполяции” — относятся к распространению результатов выборочного обследования на всю изучаемую совокупность. Ее подгруппы включают: 4) ошибки охвата и смежные ошибки, например, в отборе выборки; 5) ошибки, вытекающие из отказа от опроса или неполучения ответа; 6) ошибки выборки. Ошибки охвата чаще всего проявляются в виде систематических смещений в результате неполного охвата. Целесообразно различать неполучение ответов по причине отказа, отсутствия на месте и т.д. Подгруппы 1 — 5 вместе именуются ошибками, не связанными с отбором, в отличие от ошибок выборки (6). Последние можно подразделить на смещение отбора или оценок (которые в должным образом запланированных выборках разумного объема обычно незначительны) и дисперсию отбора.

Ошибки отбора и случайные ошибки, возникшие на стадиях сбора и обработки данных, в основном относятся к ошибкам изменчивости оценок. Они вносят в оценки элемент неопределенности, которую можно обычно выявить в рамках самого обследования: путем сопоставления результатов того, что можно назвать независимыми

внутренними повторениями опыта (т.е. на основе различных подвыборок, каждая из которых дает действительную оценку) или сопоставления независимых повторений обследования при тех же условиях.

Т а б л и ц а 20 . Виды ошибок в данных выборочного обследования

Ошибки наблюдения	1) Концептуальные ошибки 2) Ошибки при сборе данных и в ответах 3) Ошибки обработки	Смещение ответов Колеблемость ответов Ошибки редактирования Ошибки кодирования Ошибки ввода данных Неполный охват	Ошибки, не связанные с выборкой
Ошибки экстраполяции	4) Ошибки охвата 5) Ошибки неполучения ответов 6) Ошибки выборки	Излишний охват Ошибки отбора Отказ Отсутствие на месте Прочие причины Смещение оценок Колеблемость выборки	

Методы оценки качества данных

Индикаторы, или меры качества данных обследования можно получить различными путями. Некоторые методы способны дать количественную оценку масштабов и влияния различных видов ошибок, тогда как с помощью других процедур можно получить лишь качественные индикаторы. Некоторые методы подходят для получения информации об изменчивости, присущей результатам, основанным на выборочном обследовании (например, колеблемость выборки или ответов), тогда как другие методы ориентированы на оценку размеров смещений, т.е. разниц между ожидаемыми выборочными значениями и подлинными значениями интересующих нас показателей. Пригодность метода зависит от конкретного источника и вида ошибки, но следует учитывать, что отдельные этапы обследования тесно связаны между собой, поэтому ошибки не всегда можно отнести за счет конкретного их источника. Один или ряд схожих методов оценки и контроля вполне могут быть применены для измерения более чем одного вида ошибок, но некоторые из полученных индикаторов, возможно, отразят лишь общую картину точности данных, но не влияние отдельных конкретных источников и видов ошибок. Для использования и интерпретации данных необ-

ходим показатель общей точности и влияния ошибок на результаты. Но для улучшения планирования и процедур обследования следует выявить конкретные источники и элементы ошибок. Названные две цели не всегда совместимы. И происходит это по двум основным причинам. Во-первых, на практике часто можно измерить лишь некоторые, но не все ошибки. Во-вторых, измеренные таким образом элементы ошибок не всегда можно полностью отделить один от другого и они не всегда аддитивны, т.е. в совокупности они не дают общую величину ошибки. К примеру, как будет показано ниже в настоящей главе, влияние неполного охвата и неполучения ответов обычно трудноразличимы; это же относится к колеблемости выборки и ответов. Даже если теоретически разделить такого рода последствия можно, то на практике это может оказаться слишком дорогостоящим.

В зависимости от источников информации и особенностей проводимой работы различают три группы методов оценки качества данных обследования:

- 1) методы, основанные на анализе информации, полученной в ходе самого обследования, как части нормальной работы;
- 2) методы, зависящие от повторного опроса и других дополнительных операций, добавленных к основным работам обследования;
- 3) методы, основанные на сопоставлении с внешними данными из независимых источников.

Анализ информации в ходе работ по обследованию. Важно отметить, что процедуры оценки качества никоим образом не сводятся к проведению самостоятельных, сложных и общенациональных специальных исследований. Напротив, значительный объем качественной и даже количественной информации, полезной для улучшения планирования обследований, можно получить путем должной организации и документирования проводимых работ, например, предварительного опробования материалов, заслушивания отчетов интервьюеров, проверки их работы и информирования персонала о ее результатах, своевременного анализа полученных анкет и непротиворечивости таблиц. Кроме того, при минимальных дополнительных затратах и с помощью рутинных административных процедур можно получить ценную количественную информацию о многих параметрах плана обследования (например, время и расходы, связанные с поездками между выборочными районами и внутри их, повторным посещением при неполучении ответов, контролем и анализом анкет, кодированием отдельных данных, редактированием и корректировкой и т.д.). Материалы обследования можно использовать для получения таких ценных показателей, как доля

неполученных ответов и ответов, представленных за третьих лиц. При непрерывном обследовании удастся отследить динамику этих показателей.

Содержательный анализ внутренней непротиворечивости и взаимосвязей данных обследования. Весьма ценные указания относительно достоверности результатов обследования, и особенно о наличии серьезных смещений в ответах и т.д., можно получить на основе анализа итогов при их сопоставлении с уже известными знаниями и соотношениями, выявленными при аналогичных обстоятельствах, а также путем проверки на внутреннюю непротиворечивость ответов одного респондента на разные вопросы. Например, выявленные неожиданные сезонные колебания в результатах обследования могут означать, что сезонная активность не была учтена должным образом. Точно так же более низкая, чем ожидалось, степень вовлечения женщин в состав рабочей силы может указывать на то, что не удалось в полной мере выявить неоплачиваемых членов семьи и другие формы экономической активности женщин. Такой анализ, возможно, и недостаточно точен для того, чтобы выявить конкретные недостатки и ошибки в данных, но он способен по меньшей мере предостеречь пользователя от некритического восприятия результатов.

Сопоставление внутренних "повторений опыта" при обследовании. Хорошо запланированную выборку обычно можно разделить на более или менее независимые повторы или взаимопроникающие подвыборки, дающие равноценные результаты. Колеблемость оценок при этих подвыборках можно использовать для получения показателей ошибок отбора, ответа и других ошибок изменчивости оценок. Понятие взаимопроникающих выборок пояснено в парагр. 5 гл. 11, где содержится ссылка на работу (Lahiri, 1957).

Повторные опросы и другие дополнительные операции. Основная идея здесь заключается в повторении некоторой операции обследования, обычно на основе подвыборки, либо с использованием особых процедур, рассчитанных на получение более точных данных, либо как повторение работы, по сути дела, при тех же условиях.

Если усовершенствованные процедуры позволяют получать ответы, свободные от ошибок, сопоставления могут указать на конкретные ошибки на индивидуальном уровне и смещения в оценках, агрегированных на основе первоначальных данных. Примерами могут служить: повторное составление списка подмножества выборочных районов при тщательном наблюдении с целью выявления ошибки охвата; повторный опрос с тщательным уточнением для выявления ошибок в ответах; повторное кодирование специально

обученным персоналом для обнаружения ошибок кодирования. Эти и другие примеры будут рассмотрены в последующих параграфах.

Повторение при неизменных, в сущности, условиях позволяет получить оценки элементов колеблемости, т.е. изменчивости или ненадежности, присущей результатам обследования.

Сопоставление с внешними данными. В принципе, результаты обследования рабочей силы можно сопоставить с данными из внешних источников: а) на микроуровне путем сравнения записей для отдельных респондентов; б) на макро- или агрегированном уровне, где можно выявить расхождения в общей картине распределения результатов.

Сравнения на микроуровне, как правило, затруднительны, дорогостоящи и не получили распространения, но, когда они возможны, могут дать подробную информацию о характере и источниках ошибок. Сопоставления на макроуровне позволяют обнаружить лишь чистые последствия ошибок, но, когда они возможны, их осуществление не требует больших усилий и затрат. Наиболее серьезное сопоставление агрегированных данных заключается в проверке охвата выборкой и полученных оценок распределений с распределениями, выявленными в ходе переписей населения и последних оценок численности населения.

Общая стратегия оценки данных

Качество данных обследования определяется по трем характеристикам: соответствие запросам пользователей, своевременность и точность.

Соответствие запросам — это самый принципиальный и сложный аспект качества данных. Содержание и методы обследования следует строить так, чтобы были учтены показатели, которых ждут пользователи. В связи с разнообразием пользователей и вариантов применений данных обследования не всегда возможно предложить один (единственный) критерий для оценки соответствия запросам. Однако часто можно определить ситуацию, когда обследование *не сумело*, совсем или отчасти, дать информацию, отвечающую целям, ради которых оно проводилось. Вот пример. Как отмечалось выше (парагр. 3 гл. 3), имеются различные “модели” обследований рабочей силы, которые существенно отличаются по своим целям, а следовательно, содержанию и организации. Основное требование к обследованию рабочей силы — выбранная модель должна в наибольшей степени соответствовать условиям страны и обеспечивать высокое качество данных. Анализ практики разных стран показывает,

что это главное требование выполняется полностью далеко не всегда, очевидно, в результате недостаточного взаимодействия пользователей и разработчиков, отсутствия опыта и должной помощи.

Своевременность относится как к темпам работ, так и к пунктуальности: темпы определяют сроки завершения отдельных этапов работ; пунктуальность означает выдачу результатов обследования в соответствии с заранее установленным графиком. При непрерывных обследованиях, а также других обследованиях, предпринятых с вполне конкретными целями, пунктуальность может быть решающим соображением. Например, от многих обследований рабочей силы закон или сложившаяся практика требуют опубликования результатов каждого раунда к определенному сроку. При проведении более сложных, случайных, или разовых, обследований важное требование заключается в минимизации времени, необходимого для завершения обследования и выдачи его результатов. Полезность (соответствие запросам) обследования со временем убывает. Во многих развитых и некоторых развивающихся странах ежемесячные обследования рабочей силы призваны дать краткосрочные экономические показатели, и их результаты следует выдавать очень оперативно, так как польза от них быстро убывает. Если же результаты обследования рабочей силы используют для среднесрочного и долгосрочного планирования (и обследования были соответствующим образом запланированы), их ценность может сохраняться намного дольше.

Остальная часть настоящей главы посвящена третьему аспекту качества данных, а именно, *точности*. Под точностью оценок обследования обычно понимают близость к "истинным" значениям показателя для всей совокупности, которую обследование должно представлять. Как указано выше, полученным при обследовании оценкам присущи случайные ошибки, а также смещения более систематического характера. Термин *надежность* указывает, насколько малы случайные ошибки; высокая надежность означает, что оценки, полученные при повторении опыта или обследования, в общем близки друг другу, независимо от смещения, которое может быть общим для всех оценок. Высокая точность предполагает, что как переменные ошибки, так и чистое смещение малы.

Необходим также баланс между скоростью (своевременностью) и точностью (и, разумеется, между точностью различных показателей, например, текущих уровней и динамики). Пользователи данных обследований часто требуют оперативности в ущерб точности: запаздывание в получении результатов вполне очевидно, тогда как недостаточную точность результатов пользователям заметить го-

раздо труднее. Напротив, некоторые пользователи, имеющие методологическую подготовку и вовлеченные в организацию обследования, уделяют непропорционально много внимания точности в ущерб своевременности. Следует найти разумный баланс между этими двумя требованиями.

Некоторые основные принципы выбора должной стратегии.

Оценка данных

С учетом обширного опыта можно рекомендовать следующие основные принципы выбора стратегии оценки данных при обследованиях рабочей силы и других крупномасштабных обследованиях домашних хозяйств.

1) Как сказано в Резолюции 1 13-й МКСТ (ILO, 1983), важный принцип состоит в том, что оценка методологии и результатов должна всегда с самого начала рассматриваться в качестве неотъемлемой составной части планирования, организации и проведения обследования. Было бы ошибкой считать сбор большего объема данных в качестве главной задачи, а их оценку — второстепенным делом. Значимость этого принципа возрастает прямо пропорционально масштабам и охвату обследования. Он приобретает решающее значение при обследованиях, охватывающих несколько проблем; при непрерывных обследованиях с повторяющимися раундами; при крупномасштабных обследованиях; при обследованиях, ориентированных на создание банка данных многоцелевого назначения. Сказанное как раз и относится ко многим общенациональным обследованиям рабочей силы.

2) При оценке программы следует уделить внимание всем аспектам качества данных: соответствию содержания конкретным целям и области применения, своевременности (темпам и пунктуальности) и, разумеется, точности. Необходимо четко представлять себе взаимодополняющий, равно как и противоречивый характер этих требований. Поэтому требуется полный баланс между ними, особенно между своевременностью и точностью.

3) Для улучшения планирования и процедур обследования информацию о качестве данных (особенно точности) следует дополнить определенной информацией об издержках, по меньшей мере о *сравнительных* издержках при альтернативных методологиях и процедурах. В идеальном случае такую информацию надо иметь для каждого отдельного шага сбора и обработки данных. Помимо чисто материальных затрат, "издержки" должны отражать нагрузку, связанную с организацией обследования (затраты времени, по-

требность в кадрах с учетом другого возможного их использования), а также нагрузку на респондентов, предоставляющих информацию.

4) Должная оценка качества данных предполагает расходование материальных и людских ресурсов и часто требует высокой специальной подготовки. Важно выработать экономическую стратегию и ставить разумные, достижимые цели. Поэтому:

а) как правило, желательно выявить наиболее важные проблемы и на них *сконцентрировать ресурсы*, а не распылять усилия и средства. Во многих случаях анализ прошлого опыта подсказывает наиболее критические моменты, требующие улучшения: это могут быть недостатки в охвате и качестве основы выборки; слишком высокая и растущая доля неполученных ответов; неудовлетворительный объем и распределение выборки, что не позволяет делать обоснованные выводы; крупные и непредсказуемые ошибки в ответах; неудовлетворительный контроль качества и задержки в обработке данных и т.п. (этот вопрос удачно рассмотрен на основе обследования рабочей силы в Канаде в работе Fellegi and Sunter, 1974);

б) другой аспект концентрации усилий предполагает отбор *как можно меньшего числа операций* для выявления и контроля различных видов ошибок. К примеру, для выявления и контроля ошибок охвата, отбора единиц в ответах и при сверке на местах можно прибегнуть к единой всесторонней "программе наблюдения" при участии контролеров-бригадиров, вместо того чтобы иметь разные программы для каждого вида ошибок;

в) если же предпринимаются особые операции, например, повторный опрос или повторное обследование, следует четко определить их цели. Одна из наиболее частых причин неудачи таких операций с точки зрения получения нужной информации состоит в том, что они слишком амбициозны, а их цели — чрезмерно разнообразны;

г) часто в смежных обследованиях и раундах продолжающегося обследования преобладают схожие ошибки, что дает возможность со временем накапливать информацию об ошибках. С точки зрения экономии усилий еще важнее то, что показатели качества данных часто можно переносить с одной ситуации на другую. Прежде всего это относится к информации об ошибках выборки (парагр. 4);

5) Важно также уделять внимание представлению информации об ошибках обследования и отчетности в ней. Следует проводить различие между категориями пользователей информации:

а) первая категория — это общий пользователь, который скорее всего не владеет методологией обследований и содержательных исследований, а заинтересован в их результатах для получения выводов и принятия решений. Для этой группы пользователей в информации об ошибках следует отразить общее качество результатов, их недостатки и взаимосвязь с другими источниками смежной статистической информации. Говоря более конкретно, необходимо указать, каким образом неопределенность, связанная с различными источниками ошибок и смещений, способна повлиять на существенные выводы, сделанные на основе обследования. Пользователям следует не только дать информацию о качестве данных, но и сориентировать их в том, как применить результаты обследования с учетом имеющихся погрешностей;

б) вторая категория — это аналитик, занятый первичным и вторичным анализом, а также обобщением результатов. Эта группа пользователей нуждается в более детальной информации о влиянии ошибок на широкий спектр данных, полученных или потенциально возможных к получению на основе обследования. Например, такого пользователя заинтересуют не только непосредственные оценки ошибки выборки по всем основным данным, но и величина погрешности, которую можно ждать применительно к *любо*му показателю, полученному на базе обследования, по оценкам в клетках любой таблицы результатов;

в) третья категория — это статистик-организатор обследования, которого интересуют издержки и эффективность организации работ и процедур, принятых в обследовании, а также возможности изменения будущих раундов обследования. Эта группа пользователей нуждается в детальной информации об источниках и составных элементах ошибок и о том, как они соотносятся с планом, процедурами и издержками обследования.

Общему пользователю информацию об ошибках следует преподнести а) выборочно, когда она имеет значение, без ненужных деталей, особенно методологических, б) вместе с содержательными результатами обследования, а не отдельно от них. Эта информация не должна перегружать подачу содержательных результатов; цель ее предоставления — показать пределы надежности результатов, а не затушевать их.

Некоторые организации выпускают специальные отчеты о методологии и качестве данных проводимых обследований рабочей силы. Например, Статистическое бюро Канады выпускает квартальный обзор качества данных в связи с оценками в рамках обследования рабочей силы (Canada, 1982); см. также всестороннюю “картину

ошибок” в показателе занятости, учтенном в текущем обследовании населения в США (United States, 1978). Эти отчеты адресованы, очевидно, главным образом второй категории пользователей, охарактеризованных выше. Разработчики обследований могут воспользоваться более обширной информацией в виде неопубликованных отчетов о проведении обследований и их оценках на уровне отдельных опросов.

Полезные источники

Учитывая разнообразие причин и видов ошибок при обследованиях рабочей силы и других обследованиях домашних хозяйств, по этому вопросу издают обширную литературу. Например, опубликован всесторонний анализ причин, оценки и контроля ошибок, не связанных с выборкой, при проведении обследований домашних хозяйств (United Nations, 1982). Помимо подробного описания различных источников и видов ошибок в этой работе содержатся обширная библиография и много примеров ошибок в выборочных обследованиях и переписях из опыта развитых и развивающихся стран. Развернутая библиография приводится также Далениусом (1977 г.). “Классической” работой об измерении ошибок обследования является статья Хансена и др. (1961 г.). Краткое и четкое описание методов оценки ошибок в переписях и обследованиях содержится в работе, относящейся к Канаде (Canada, 1978). Различные виды ошибок и анализ основных вопросов их оценки применительно к обследованиям домашних хозяйств в развивающихся странах приведены в работе (Verma, 1981).

2. ОШИБКИ ОХВАТА И СМЕЖНЫЕ ОШИБКИ

Для того, чтобы результаты выборочного обследования можно было распространить на всю изучаемую совокупность, необходимо, чтобы каждый элемент совокупности имел некоторую ненулевую вероятность быть отобранным для выборки (это именуется “вероятностным отбором”, см. парагр. 4 гл. 11). Указанное условие нарушается, если: а) изучаемая совокупность полностью не представлена в основе выборки; б) отбор единиц из основы не произведен согласно процедурам, предусмотренным планом выборки; в) не все отобранные единицы выборки были должным образом обследованы. В этом разделе рассмотрены ошибки охвата (а) и смежные ошибки отбора (б); ошибки, относящиеся к неполучению ответов (с), рассмотрены в парагр. 3.

Характеристика ошибок охвата

На практике выборку отбирают из основы, представляющей собой список, другую требуемую вспомогательную информацию, а также определенный набор правил и процедур увязки единиц различного вида. Ошибки охвата возникают, когда:

1) некоторые единицы обследуемой совокупности отсутствуют в основе для отбора. Это — “неполный охват”: отсутствующие единицы не имеют шанса попасть в число отобранных из основы;

2) некоторые единицы в основе не принадлежат к изучаемой совокупности. Это ведет к “излишнему охвату”, если эти единицы, будучи отобранными, не могут быть идентифицированы и исключены на этапе реализации выборки.

3) Некоторые единицы обследуемой совокупности появляются в основе более одного раза (“дублирование”), что увеличивает их шанс быть отобранным для выборки.

Указанные проблемы могут возникнуть в явном виде при использовании некачественных списков или в неявном виде в результате путаницы в границах единиц или правилах увязки единиц различного вида. При обследованиях домашних хозяйств основной проблемой обычно является неполный, а не излишний охват.

Во многих обследованиях рабочей силы, как и других крупномасштабных обследованиях домашних хозяйств, отбор носит многоступенчатый характер и, как правило, включает:

отбор территориальных единиц с помощью одной или нескольких ступеней;

составление списков и отбор зданий, жилых единиц или домашних хозяйств в пределах отобранной самой низкой ступени территориальных единиц;

составление списка отдельных индивидов в пределах отобранных домашних хозяйств и отбор индивидов с определенными характеристиками для обследования.

Ошибки охвата могут возникнуть на любой из этих ступеней.

Территориальные единицы

В большинстве случаев можно рассчитывать на полноту списков территориальных единиц. Однако серьезные проблемы охвата могут возникнуть в связи с установлением *границ* территориальных единиц, которые были бы четкими, исчерпывающими и непересекающимися. Приведем несколько примеров:

а) во многих ситуациях имеющаяся основа территориальных единиц — это фактически список населенных пунктов, например, де-

ревень, а не действительно территориальных единиц, охватывающих всю обследуемую совокупность. Обычно это обусловлено недостатками картографической работы при переписях, которая создает основу для обследований домашних хозяйств. Обособленно расположенные домашние хозяйства и лица, не проживающие в каком-либо населенном пункте, вполне могут не попасть в такую основу;

б) новые поселения нередко не учтены в устаревшей основе;

в) схожая проблема заключается в изменении названий и границ административных единиц; при наличии таких изменений отобранные единицы и единицы, которые значатся в основе, могут не совпадать;

г) из-за неудовлетворительного качества карт и описаний персонал обследования часто не может установить и четко разграничить отобранные районы. Такие ошибки идентификации носят систематический, а не случайный характер;

д) во многих развивающихся странах наименьшие единицы, имеющиеся для целей отбора, слишком велики, чтобы служить в качестве эффективной низшей ступени отбора территориальных единиц. Для членения имеющихся единиц используются специальные методы. При этом может быть привнесен элемент систематических смещений, особенно когда членение поручают самим интервьюерам без должного контроля за ними. Было установлено, например, что при некоторых обследованиях в Индии счетчики склонны были отбирать небольшие участки, очевидно, чтобы сократить нагрузку при последующем опросе. Подобное может привести к существенно неполному охвату. Систематические смещения возникают и тогда, когда работникам предлагают создавать участки заранее определенной одинаковой величины.

Жилые единицы и домашние хозяйства

Списки жилых единиц (домов) менее стабильны, чем основы территориальных единиц, а списки домашних хозяйств обычно менее стабильны, нежели списки жилых единиц или домов. Главная проблема на ступени отбора жилых единиц или домашних хозяйств — это неполный охват, который возможен в результате: а) неполноты и низкого качества списков в пределах отобранных территориальных единиц; б) разрывы во времени между составлением списков и опросом, что способно привести к тому, что вновь возникшие жилые единицы и домашние хозяйства не найдут отражения в основе выборки. Охват иногда можно улучшить путем составления списков и отбора единиц в пределах соседних групп, а не поодиночке. Можно

также составить дополнительную выборку вновь построенных жилых единиц, если такая информация имеется в органах, выдающих соответствующие разрешения. *Излишний охват* также возможен (хотя на практике он встречается реже, чем неполный охват), когда: а) некоторые единицы попали в список более одного раза; б) включены единицы вне рамок обследования, и они не были выявлены и исключены на этапе сбора данных; в) домашние хозяйства и жилые единицы из территорий, не отобранных для выборки, включены в списки для отобранных территориальных единиц. Необходимо уделять большое внимание и предусматривать достаточные ресурсы для составления и ведения обновляемых списков с целью выявления ошибок, связанных с неполным или излишним охватом.

Отбор отдельных лиц

При обследовании рабочей силы значительная часть собранной информации относится к отдельным лицам, которых выявляют и заносят в списки в рамках домашних хозяйств. Индивиды подлежат опросу в ходе обследования на основе определенных индивидуальных характеристик, например возраста. Неполный охват на этом этапе может произойти в результате: а) неспособности выявить некоторых лиц, которых следовало бы опросить в отобранном домашнем хозяйстве; б) неправильной информации об индивидуальных характеристиках, требуемых для отбора. Наиболее распространенными категориями неучтенных лиц являются юноши, жильцы, прислуга и другие члены домашнего хозяйства, не являющиеся членами семьи, лица, занимающиеся незаконной деятельностью или деятельностью с помощью незаконных средств. При охвате на основе определения “де-факто” (см. парагр. 2 гл. 11) лица “в пути” (но обычно проживающие в данном домашнем хозяйстве) не опрашиваются.

Ошибки отбора и реализации выборки

Аналогичные ошибки могут возникать из-за неправильного применения процедуры отбора. Эти ошибки отличаются от ошибок охвата в строгом смысле слова тем, что последние относятся к дефектам основы, ее неполноте, тогда как ошибки отбора связаны с самим процессом получения выборочной совокупности. Поэтому методы выявления и контроля применительно к этим видам ошибок, как правило, различны. Однако последствия их, в сущности, одинаковы: некоторые единицы совершенно исключаются из отбора (т.е.

вероятность их включения в выборку равна нулю), тогда как другие отбираются с вероятностью, отличной от плановой, и обычно это нигде не фиксируется. Эти ошибки ведут к искажению структуры выборки, что способно привести к известным и неизвестным смещениям в статистических оценках, полученных в результате обследования.

Некоторые ошибки отбора, если они выявлены, можно отчасти скорректировать на стадии оценки путем присвоения соответствующих весов выборочным единицам. Но это не всегда возможно, особенно когда процедуры отбора должным образом не регистрируются.

На этапе реализации выборки систематические ошибки охвата могут быть результатом подмены: например, когда интервьюер заменяет труднодоступную единицу, отобранную для опроса, более доступной единицей.

Последствия ошибок охвата и смежных ошибок

Помимо величины и характера данной ошибки ее последствия зависят от вида затронутого ею показателя. При оценке совокупных показателей для всего населения, например общего числа безработных, последствия ошибки охвата проявляются непосредственно и имеют ту же относительную величину. Иначе говоря, доля ошибки в показателе для всего населения будет такой же, как и в выборке с неполным (реже — излишним) охватом (если, разумеется, не была осуществлена корректировка на этапе оценки с использованием более надежных показателей из некоторых внешних источников). При оценке показателей типа общего количества отработанных часов последствия будут зависеть не только от числа пропущенных единиц, но и от их относительной значимости, т.е. от того, работали ли пропущенные единицы чаще больше или меньше среднего количества часов. Так, определенные сложности возникают при некоторых обследованиях сельскохозяйственных и других предприятий, где показатели различаются, причем на небольшое число полностью охваченных единиц приходится значительная доля общего показателя. Но при обследованиях рабочей силы, когда единицами выступают домашние хозяйства и индивиды, значения многих показателей для отдельных единиц скорее схожи, и влияние неполного охвата на всю совокупность примерно пропорционально величине самого неполного охвата. Но это может быть не всегда так: чаще оказываются неохваченными небольшие домашние хозяйства или отдельно живущие лица, которые включают экономически актив-

ных индивидов, занятых на работе, и поэтому часто не попадающих в обследование.

Что же касается оценок *долей, средних, уровней и других коэффициентов* (в отличие от оценок абсолютных величин для всей совокупности), то смещения, вызванные неполным охватом, в каком-то смысле выступают не так прямо и зависят от средних характеристик единиц, охваченных и не охваченных обследованием. Эти последствия существенны, если различия в характеристиках и степень неполноты охвата значительны.

Смещения, связанные с ошибками охвата, при оценке *различий* между отдельными подгруппами населения или периодами времени (например, сельским и городским населением, сезонными уровнями безработицы) зависят от чистой (алгебраической) разности между смещениями для рассматриваемых подгрупп или периодов. Однонаправленные смещения в тенденции погашают друг друга, тогда как разнонаправленные смещения аддитивны; к счастью, первый вариант встречается чаще, но он не единственный. Обратите внимание, что смещение, незначительное при его сравнении со средней для группы (или периода), может быть значительным, если его сравнивать с разностью между средними по группам (или периодам).

Оценка последствий ошибок

Ни размеры, ни последствия ошибок охвата не поддаются легкой оценке, ибо она требует наличия информации внешней не только по отношению к выборке, но также и, по определению, к основе выборки.

Макросопоставления с внешними данными

Основное требование к любому выборочному обследованию заключается в необходимости удостовериться, что полученная выборка соответствует по распределению и характеристикам изучаемой совокупности, которую она призвана отразить (если не считать изменчивости, связанной с ошибкой отбора). Наиболее надежный путь проверки результатов выборочного обследования рабочей силы — это сопоставить их с текущими оценками населения на основе переписей, актов регистрации и других источников. Важно, чтобы источники, с которыми проводят сравнения, были сопоставимы по охвату и содержанию, как в территориальном, так и во временном отношениях. Например, лица, живущие вне частного сектора домашних хозяйств или в отдаленных районах, а также некоторые мобильные индивиды и т.д. обычно охвачены переписями, но откры-

то исключаются из обследования. Естественно, что при сравнениях необходимо сделать поправку на указанные различия.

В идеальном случае следует сравнивать оценки для всей совокупности на основе выборки с оценками на основе внешнего источника, причем те и другие оценки группируют по одним и тем же признакам — географическое расположение, тип населенного пункта, пол, возраст и другие характеристики населения. Показатели полноты и представительности выборки получают путем исчисления следующего коэффициента для различных подгрупп:

$$c = \frac{\text{оценка для всей подгруппы на основе выборки}}{\text{та же оценка на основе надежного и сопоставимого источника}}$$

Обычно проблема заключается в неполном охвате и c тогда меньше 1,0. Величину s , равную $1 - c$ именуют *коэффициентом сдвига* (slippage rate).

Такое сравнение дает лишь общее представление о чистом охвате, т.е. чистом следствии ошибок неполного или излишнего охвата. Оно не отражает валовую ошибку охвата (т.е. общее число ошибок охвата — как неполного, так и излишнего), и, как правило, неспособно подсказать источник ошибки. Однако иногда представляется возможным предположить источники расхождений, а затем исследовать их, если это практически целесообразно, с помощью более точных методов. Фактически при таком сравнении получают представление об общей представительности полученной выборки (а не о чистом охвате как таковом), в чем находит отражение, помимо полноты основы, влияние других источников ошибок, например, ошибок отбора и реализации, неполучения ответов, ошибочных ответов и неизбежной изменчивости отбора. Однако во многих ситуациях преобладают следствия ошибок в чистом охвате (а также отбора и реализации выборки).

Если исключить последствия общей чистой ошибки охвата, искажения в распределении выборочной совокупности можно выявить, сопоставив *взвешенное выборочное распределение с распределением генеральной совокупности* при тех же группировочных признаках:

$$d = \frac{\text{доля подгруппы населения во взвешенной выборке}}{\text{доля этой же подгруппы на основе внешнего источника}}$$

Коэффициент c позволяет оценить итоговые показатели для всей совокупности на базе выборки, тогда как d , основанный на менее жестких требованиях, относится к оценкам средних, долей и других коэффициентов на базе выборки.

Как пояснено в парагр 6 гл. 11, обратная величина коэффициента (c или d) представляет собой основные веса, которые можно применить к выборочным данным (уже взвешенным по величинам, обратным выборочным долям) с целью связки контрольных итогов и долей с внешними показателями. Уже отмечалось, что, делая такие поправки, важно убедиться, что внешние показатели действительно точнее оценок и долей, полученных непосредственно на базе выборки; в противном случае можно получить большие и неожиданные смещения. Указанное требование выполняется далеко не всегда, как это происходит во многих развивающихся странах, где внешние данные о населении непригодны для сопоставлений и недостаточно надежны. В частности, опасность появления смещения особенно велика, если поправочные веса применяют отдельно для многих подгрупп при слишком дробной классификации, например, для небольших регионов или малочисленных категорий населения.

Сравнения на микроуровне

Более точное представление о валовой ошибке охвата, а также некоторую информацию о возможных источниках ошибок можно получить путем тщательного сравнения списков обследования с имеющимися списками из других источников. К числу последних относятся: а) внешний источник типа реестра населения или списков избирателей; б) специальная операция по пересоставлению списков домашних хозяйств и/или членов этих домашних хозяйств с помощью какого-либо более совершенного метода; затем осуществляют сравнение на микроуровне.

Вариант (а) находит ограниченное применение в развивающихся странах, так как здесь, как правило, отсутствуют надежные списки домашних хозяйств и индивидов. Где этот метод применим, он может дать хорошие оценки неполного охвата в той мере, в какой независимый источник информации полон, сопоставим во времени и пространстве и содержит достаточную информацию для сравнения с единицами основы выборки на микроуровне. А так как этот метод ориентирован на выявление индивидуальных единиц, не охваченных основой выборки, он может потенциально пролить свет на причины неполного охвата и характеристики неучтенных единиц. Недостатки метода таковы: трудности в получении надежных, сопоставимых и независимых внешних данных; сложность, ошибки и высокие затраты, связанные с детальным сравнением списков.

Вариант (б) требует отдельной работы по пересоставлению списков домашних хозяйств и/или членов, например, в рамках по-

вторного обследования или опроса. По сравнению с (а) этот метод может дать более сопоставимую информацию, которую легче также сравнить со списками обследования. Главный недостаток этого метода — высокие затраты по проведению дополнительной специальной работы. Поэтому ее, как правило, можно провести в ограниченных рамках, например, по небольшому числу выборочных районов. При этом территориальные единицы следует отобрать таким образом, чтобы результаты можно было распространить на всю изучаемую совокупность с определенной степенью достоверности. Целесообразно включить районы, где ожидаемый уровень ошибок выше среднего. В то же время выборка должна быть достаточно мала и компактна, что обеспечит тщательное наблюдение и контроль. Основное достоинство варианта (б) состоит в том, что при отсутствии во многих случаях внешних источников это единственно практически возможный метод изучения валовых ошибок охвата.

3. НЕПОЛУЧЕНИЕ ОТВЕТОВ

Неполучение ответов означает неспособность добиться требуемой информации: а) от единиц, включенных в выборку, б) по некоторым вопросам анкеты, относящимся к этим единицам. Отсутствие информации по отдельным единицам выборки именуют “полным неполучением ответов”. Приведенная градация не является очень строгой. Например, при непрерывном обследовании рабочей силы некоторые единицы остаются в выборке в ходе нескольких раундов, и часто можно добавить некоторую информацию (например, о возрасте, поле и т.д.) из более раннего раунда в случае, когда не получен ответ. Эта процедура приводит к тому, что полное неполучение превращается в частичное неполучение ответов. Напротив, недостающая информация в ряде случаев частичного неполучения ответов может быть настолько существенной с точки зрения целей обследования, что мало смысла сохранять частично полученную информацию, и случаи частичного неполучения превращаются таким образом в случаи полного неполучения ответов.

Тем не менее выделение полного и частичного неполучения ответов полезно. Первое характеризует обследование в целом: как оно воспринимается, насколько оно сложно, как организован и контролируется сбор данных. Второе отражает сложность и ясность отдельных пунктов информации в анкете, а также степень понимания и прилежания отдельных интервьюеров. По своим источникам, влиянию и методам оценки и контроля частичное неполучение ответов схоже с ошибками в ответах, о которых говорится в парагр. 5. В этом

параграфе в основном рассмотрены последствия полного неполучения ответов (которое ниже именуется просто “неполучение ответов”); этот фактор обычно учитывают на уровне отдельных *счетных единиц*, которые могут совпадать и не совпадать с единицами самой последней ступени отбора. Когда же неполучение учитывают на уровне отдельных пунктов информации, это, естественно, характеризует последствия полного и неполного получения ответов на этот вопрос.

Источники

Неполучение ответов может быть обусловлено различными факторами, и их влияние неодинаково в различных странах и в разных обследованиях. Во многих развитых странах, особенно в крупных городах, в некоторых развивающихся странах проблема, похоже, все больше заключается в сознательном отказе отвечать или в отсутствии респондентов дома в момент прихода интервьюера. Но в целом, и особенно в сельских районах многих развивающихся стран, прямые отказы все еще относительно редки. Более характерная проблема — это трудности вступления в контакт и идентификации выборочных единиц, что в свою очередь обусловлено изъятиями в информации, содержащейся в основе выборки, а также другими трудностями, например физической недоступностью. В ряде стран закон требует обязательного участия в официальных переписях и обследованиях. И хотя законодательство, если его удастся провести в жизнь (а это не всегда так), способно увеличить долю полученных ответов, оно может также привести к серьезным ошибкам в ответах, когда отдельные респонденты не склонны добровольно участвовать в таких мероприятиях.

Последствия

Последствия неполучения ответов для итогов обследования, в принципе, такие же, как и последствия неполного охвата. Если невозможно сделать поправки на стадии оценки, то неполучение ответов оказывает прямое влияние на оценку итогов для всей совокупности: доля ошибки в оценке общего количества единиц в совокупности близка к доле неполученных ответов. При оценке общих *показателей* последствия также будут зависеть от относительных показателей отсутствующих единиц. Смещения в оценках средних, долей и других коэффициентов возникают лишь в той мере, в какой неответившие единицы отличаются по своим харак-

теристикам от ответивших единиц; такие систематические расхождения между этими видами единиц, разумеется, встречаются довольно часто.

Помимо добавления недостающих показателей, когда это целесообразно и возможно, поправки на неполучение ответов обычно делают путем увеличения выборочных оценок на величину, обратную значению *коэффициента ответов* (см. ниже). Для указанной цели выборку обычно делят на несколько частей или слоев, которые, как полагают, более внутренне однородны, нежели совокупность в целом. Примерами могут служить деление территорий на небольшие географические районы или типы населенных пунктов; этнические и социально-демографические характеристики индивидов можно использовать, если они имеются, для классификации респондентов и нереспондентов. Затем находят поправку отдельно для каждой группы нереспондентов. В результате смещения сохраняются в оценках в той мере, в какой респонденты и нереспонденты отличаются друг от друга *в пределах* этих отдельных частей, групп или слоев.

Обобщая, можно сказать, что оценки влияния неполучения ответов на результаты обследования зависят от а) частоты, или коэффициента, неполучения ответов и б) характеристик нереспондентов в пределах категорий, на которые были разбиты обе эти группы лиц. В принципе, исчисление (а) осуществляют просто, хотя, как показано ниже, могут возникнуть двусмысленности и практические затруднения. Получить информацию относительно (б), как правило, намного сложнее, а часто невозможно. Иногда информацию об основных характеристиках нереспондентов можно получить от соседей, других членов семьи и т.д. в ходе обследования. При непрерывных обследованиях иногда можно добавить информацию из предыдущих раундов. В ряде случаев предпринимались активные меры по опросу тех, кто не дал ответов. Этот вариант рекомендуется тогда, когда доля неполученных ответов велика и способна обесценить обследование.

Влияние неполучения ответов зависит от их причины. Так, лица, отсутствующие при обследовании, — это нередко экономически активные индивиды, проживающие одни или в небольших домашних хозяйствах. Эти лица часто отличаются от респондентов в большей мере, нежели другие лица, которые не были опрошены. Следовательно, влияние этого фактора на результаты обследования будет также более значительным (см., например, Hong Kong, 1980; Sweden, 1980a, 1980b). Вывод сводится к тому, что важно регистрировать причины отдельных случаев неполучения ответов.

Подобно тому, как ошибки охвата характеризуют качество основы выборки, включая качество карт и списков, масштабы неполучения ответов могут служить наиболее красноречивым показателем качества и успешности реализации обследования. Независимо от его влияния на результаты обследования и возможности устранения или уменьшения смещений путем внесения поправок на стадиях обработки или анализа, высокий процент неполучения ответов означает, что обследование в целом было плохо воспринято и/или плохо организовано.

Измерение: исчисление коэффициента неполучения ответов

В отличие от неполного охвата, неполучение ответов обычно можно измерить на основе внутренней отчетности обследования. В принципе, коэффициент определяется просто: число подходящих единиц выборки, не давших ответа, делится на общее число единиц выборки. Чтобы воспользоваться приведенным определением, необходимо четко уяснить, что такое “подходящая” единица и что означает “давших ответ”. Ниже мы попытаемся вывести точную формулу для исчисления коэффициента неполучения ответов.

Для начала рассмотрим простой случай, когда отбирают самовзвешивающуюся выборку из некоторого числа n домашних хозяйств, из которых все подходят для обследования. Из указанного числа n' домашних хозяйств были успешно опрошены. Успешным можно назвать опрос, при котором получена информация, по меньшей мере по определенному кругу важнейших ее элементов. Коэффициент получения ответов R мы определяем так:

$$R = n'/n \text{ (обычно в процентном отношении),} \quad (1)$$

а соответствующий коэффициент неполучения ответов (NR) как:

$$NR = (1 - R) = \frac{(n - n')}{n}$$

На практике ситуация обычно сложнее представленного выше простого случая. Основная трудность связана с тем, что объем отобранной выборки (n) неизвестен, если конечные единицы отбора, перечисленные в списке последней ступени отбора, по своему типу отличаются от учтенных единиц. Это затруднение во многих обследованиях усугубляется отсутствием четкого определения списочных единиц.

Важно отметить, что числа в уравнении (1) относятся к единицам учета, а не обязательно к единицам отбора. Последние нередко представляют собой гнезда учтенных единиц. Рассмотрим обследование, для которого была отобрана выборка из m жилых единиц, и из них с m единиц был успешно найден контакт в ходе обследования. Предположим, что в этих m жилых единицах установлено наличие n домашних хозяйств, из которых n' были успешно опрошены в ходе обследования.

Чтобы исчислить коэффициент получения ответов для домашних хозяйств, мы точно знаем число успешно опрошенных (n'), но не точно — число “отобранных” единиц. Последнее число можно установить, лишь прибегнув к некоторым допущениям. На практике исчислить коэффициент можно следующим образом.

Процесс сбора данных включает два этапа: контакт с жилыми единицами выборки и опрос домашних хозяйств в этих жилых единицах. На первом этапе коэффициент получения ответов равен:

$$R = \frac{\text{число жилых ед., с которыми успешно вступили в контакт}}{\text{число отобранных жилых единиц}} = \frac{m'}{m}$$

На втором этапе он равен:

$$R = \frac{\text{число успешно опрошенных домашних хозяйств}}{\text{общее число домашних хозяйств в жилых единицах, с которыми вступили в контакт}} = \frac{n''}{n'}$$

Общий коэффициент получения ответов для опроса домашних хозяйств равен произведению:

$$R = \frac{m'}{m} \cdot \frac{n''}{n'}$$

Обратите внимание, что в жилых единицах, с которыми *вступили в контакт*, среднее число домашних хозяйств в расчете на одну жилую единицу равно n'/m' . Если предположить, что эта же средняя применима и к жилым единицам, с которыми контакт не установлен, тогда оценка числа домашних хозяйств во всех жилых единицах выборки составит:

$$n = m (n'/m'),$$

(число жилых единиц в выборке среднее число домашних хозяйств в одной единице),

при этом значении n уравнение (1) принимает вид:

$$R = \frac{n''}{n}, \text{ т.е. } \frac{\text{число опрошенных домашних хозяйств}}{\text{оценка числа отобранных домашних хозяйств}} \quad (2)$$

Очевидно, что в основе уравнения (2) лежит допущение, о котором говорилось выше: среднее число домашних хозяйств в расчете на жилую единицу одинаково в единицах, с которыми установлен и не установлен контакт.

Эти же идеи можно распространить на следующую стадию: составление списка лиц в обследуемых домашних хозяйствах и их опрос. Коэффициент ответов на третьей стадии равен:

$$R = \frac{\text{число успешно опрошенных лиц}}{\text{число лиц, подлежащих опросу в обследуемых домашних хозяйствах}} = \frac{p'''}{p''}$$

Общий коэффициент ответов на трех стадиях анализа (жилые единицы, домашние хозяйства, индивиды) составит:

$$R = (m'/m)(n''/n')(p'''/p'') = p'''/p, \text{ где } p = m(n'/m')(p''/n'') \quad (3)$$

Как и раньше, числитель (p) — это число опрошенных лиц, а знаменатель (p) — оценка количества индивидов, отобранных для выборки. Сделано еще одно допущение, а именно, среднее число подлежащих опросу лиц в одном домашнем хозяйстве одинаково в охваченных и не охваченных обследованием домашних хозяйствах. Это допущение не всегда оправдано: замечено, например, что с одинокими лицами и небольшими домашними хозяйствами установить контакт намного сложнее. Однако уравнение (3) можно применить отдельно для домашних хозяйств различного размера. Но это потребует наличия информации о размерах домашних хозяйств, которые дали и не дали ответы, или какого-либо допущения относительно распределения домашних хозяйств по размерам в отобранной выборке. Такую информацию легче получить при непрерывных обследованиях, когда выборки, учтенные в отдельных районах, пересекаются или соотносятся еще каким-либо образом.

Пробелы в выборочных списках

Трудности при исчислении коэффициентов неполучения ответов возможны в связи с наличием пробелов в выборочных списках. "Пробелы" — это позиции, представляющие собой несуществующие или выходящие за рамки обследования единицы. Эти пробелы следует игнорировать, так как неполучение ответов относится лишь к единицам из подлежащей изучению совокупности. Если пробе-

лы можно правильно выявить и исключить из выборочных листов, то проблемы с точки зрения неполучения ответов не возникает. (Разумеется, если они выявлены лишь после отбора выборки, их присутствие может изменить объем выборки и увеличить дисперсию отбора.) На практике же проблема состоит в том, что при опросе возможно спутать некоторые из таких пробелов с действительными случаями неполучения ответов. Неудовлетворительный контроль и ведение документации, а также использование устаревших списков для отбора не только ведут к увеличению случаев неполучения ответов, но и усложняют различение пробелов и неполучения ответов. Например, когда опросу домашних хозяйств предшествует отдельная операция по составлению списков жилых единиц, значительная доля неполученных ответов может состоять из единиц, которые значатся в списках и все еще существуют, но не были найдены интервьюером из-за плохой документации и связи. В то же время в списках могут содержаться несуществующие адреса (сношенные дома, ошибки в списках и т.д.). Из названных двух групп первая — это неполученные ответы, вторая — пробелы в основе выборки. На практике различить их часто бывает трудно. Точно так же интервьюеры часто неспособны различить пустующие жилые единицы (что равносильно пробелам в списках) и занятые жилые единицы, обитатели которых временно отсутствуют (что, в зависимости от определения охвата, может означать неполучение ответов). Важное требование состоит в том, чтобы с помощью четко определенных категорий интервьюер мог выделить результат контактов с каждой единицей и подлинное неполучение ответов можно было бы отличить от пробелов или единиц списка, не подлежащих опросу. Тогда пробелы можно будет полностью исключить при исчислении коэффициентов. Когда указанное требование не выполняют, возникают сомнительные случаи и можно исчислить лишь определенный диапазон, в пределах которого находится коэффициент получения ответов: верхняя оценка будет задана в предположении, что все сомнительные случаи это пробелы (т.е. подлежат исключению и в числителе и в знаменателе формулы); нижнюю оценку получают при предположении, что это только единицы, от которых не получены ответы. Эти вопросы с рядом примеров из развивающихся стран рассмотрены в работе (Verma, 1980).

Взвешенные выборки

При несамовзвешивающихся выборках вопрос заключается в том, будут ли коэффициенты получения ответов исчислены на основе взвешенных или невзвешенных данных.

Для контроля хода работ и аналогичных целей часто достаточно иметь коэффициенты на основе невзвешенных данных: в этом случае они показывают действительное количество завершенных и незавершенных опросов. Их легче исчислить и за ними проще следить в динамике, в зависимости от обследуемого района, по каждому интервьюеру и т.п.

Чтобы получить более точное представление о *последствиях* неполучения ответов для результатов обследования, предпочтительнее исчислять взвешенные коэффициенты, когда соответствующие выборочные веса применяют к единицам выборки в числителе и знаменателе уравнений типа (1) — (3). Во всяком случае, когда в результаты выборки вносят поправки, т.е. их умножают на величину, обратную коэффициенту получения ответов, существенно, чтобы коэффициенты были исчислены на основе взвешенных данных.

Некоторые примеры коэффициентов неполучения ответов при обследованиях рабочей силы

Приводимые ниже цифры, взятые из работы МБТ (1986), дают представление об уровне неполучения ответов в обследованиях рабочей силы. Выясняется, что примерно половина стран, по которым имеются сведения, могут удерживать коэффициенты неполучения ответов при проведении обследований рабочей силы в пределах 5%. Сюда входят некоторые развивающиеся страны, а также развитые страны — Австрия, Западная Германия, Италия, Канада, США и Япония. Однако часто встречаются коэффициенты 5—12% (примерно в 1/3 стран, включая Испанию, Малайзию, Норвегию, Перу и Филиппины). Некоторые страны сообщили о весьма высоких коэффициентах неполучения ответов, причем наиболее высокий из них в Дании (24%). Обескураживает, что значительная доля стран не сообщает данные о коэффициентах неполучения ответов.

Общее количество стран, охваченных исследованием МБТ		44
Количество стран:		
— не сообщивших данные о коэффициентах		14
— неполучения ответов		
— сообщивших такие данные, всего		30
из них коэффициенты равны	<2%	5
	2—5%	11
	5—8%	5
	8—12%	6
	более 12%	3

Дополнительные показатели

Помимо общих коэффициентов получения и неполучения ответов часто полезно также исчислить и контролировать некоторые дополнительные показатели. Вот ряд из них:

1) *Коэффициенты неполучения ответов в разрезе причин.* Число незаконченных опросов можно сгруппировать в зависимости от главной причины неполучения ответов. Сгруппированные таким образом неполученные ответы делят на (действительное или оцененное) число отобранных единиц и получают коэффициенты в разрезе причин, например, отказа, отсутствия контактов и т.д. В сумме эти частичные коэффициенты дают общий коэффициент неполучения ответов.

2) *Коэффициенты получения ответов по стадиям.* Обследование нередко протекает в виде стадий — от конечных элементов отбора (т.е. жилых единиц) до счетных единиц (т.е. индивидов). Полезно исчислить коэффициенты для каждой стадии, причем их произведение будет равно общему коэффициенту полученных ответов. Коэффициенты по стадиям целесообразно классифицировать в зависимости от причин неполучения ответов.

3) *Коэффициенты неполучения ответов по каждому из интервьюеров.* Как отмечалось выше, коэффициенты неполучения ответов по каждому интервьюеру служат важным инструментом контроля особенно при непрерывных обследованиях.

4) *Пробелы в выборочных листах.* Когда листы содержат много пробелов, может оказаться полезным исчисление коэффициентов наличия пробелов, возможно после их классификации по видам (например, снесенные дома, незанятые единицы, за рамками обследования и т.д.). В знаменателе указывают общее число отобранных единиц для выборки, включая пробелы в списках; в числителе — количество пробелов по видам, если это возможно.

5) *Неполучение ответов на отдельные пункты.* На этапе анализа при табулировании и истолковании результатов может представиться важным учесть отсутствующие данные по отдельным элементам. Поэтому необходимо, чтобы все добавленные значения при неполучении ответов были как-то выделены в массивах данных, с тем чтобы их можно было впоследствии выявить. Эти пункты не обязательно соответствуют отдельным вопросам анкеты, а могут быть в виде показателей, полученных на основе ряда вопросов. Обычно вычисления ограничивают случаями, в которых опрос в целом завершен успешно (возможно, за исключением некоторых пунктов, в том числе представляющих интерес) и которые содержат пункт, представляющий интерес. В знаменателе будут все эти слу-

чаи, а в числителе — те из них, в которых на пункт, представляющий интерес, ответ не был получен:

коэффициент неполучения ответов на определенный пункт $= a/b$, где

а) “завершенные” случаи, по которым не получен ответ на пункт, представляющий интерес,

б) все “завершенные” случаи, где применим пункт, представляющий интерес.

В приведенном выше выражении для исчисления коэффициента неполучения ответов на определенный пункт не учтены случаи, исключенные из-за полного отсутствия ответов. Обычно это все, что необходимо. Для того чтобы учесть влияние полного неполучения ответов (оценить полные последствия неполучения ответов на этот пункт), необходимо в числитель и знаменатель добавить одну и ту же величину. Эта величина представляет собой оценку количества случаев из числа тех, по которым полностью отсутствуют ответы и к которым имеет отношение пункт, представляющий интерес. Это количество не всегда имеется, но его можно получить на основе распределения случаев, давших ответы.

4. ОШИБКА ВЫБОРКИ

Повсеместно считается правильным, когда результаты обследования сопровождаются информацией об изменчивости оценок в связи с отбором элементов, а при истолковании результатов обследования этот фактор принимают во внимание. Сказанное особенно важно применительно к крупномасштабным непрерывным обследованиям.

Ошибки отбора присущи процессу оценки параметров совокупности на основе результатов, полученных на базе вероятностной выборки из совокупности. План выборки определяет правила, по которым следует отбирать единицы для учета и опроса и по которым оценивают параметры всей совокупности; даже при отсутствии ошибок измерения и реализации выборки, повторная реализация приводит к другим оценкам в зависимости от отобранных единиц. Ошибка выборки характеризует меру изменчивости оценки при теоретически возможных повторениях обследования и отсутствии других ошибок, не связанных с выборкой.

Значимость ошибок выборки

При истолковании информации об ошибках выборки следует помнить, что они лишь часть совокупной ошибки обследования: это

нижний предел совокупной ошибки. Поэтому информация об ошибках выборки носит “характер ориентира”. Это означает, что если ошибка выборки слишком велика, результаты обследования в целом могут оказаться бесполезными для практических целей. Необходимо (но никоим образом не достаточно), чтобы ошибки выборки были в известных пределах, ибо только тогда на основе результатов обследования можно делать существенные выводы.

Кроме того, информация о величине ошибок выборки важна для определения степени детализации, с которой можно обоснованно классифицировать данные обследования. При обследованиях рабочей силы правильное истолкование результатов обследования обычно предполагает весьма детальную группировку результатов по признаку пола и возраста, другим характеристикам индивида и домашнему хозяйству, типам населенных пунктов, географическим районам и т.д. Даже при выборке в несколько тысяч респондентов числа в клетках таблиц быстро могут стать очень небольшими. Грубо говоря, в то время как величина смещения в группе, не связанного с выборкой, не зависит от ее объема, изменчивость, связанная с отбором (а также некоторые другие ее элементы), как правило, возрастает по мере уменьшения объема выборки. Следовательно, объем вполне может стать самым существенным фактором совокупной ошибки во многих небольших группах совокупности и при представляющих интерес сопоставлениях.

Информация об ошибках выборки существенна также при составлении плана, оценке и изменении плана выборки. Величина ошибки выборки в той или иной оценке показателя зависит, помимо прочего, от объема и плана выборки, в частности, от степени гнездования единиц в выборке и их однородности внутри гнезд. Чтобы уменьшить ошибку выборки, необходимо увеличить ее объем и/или уменьшить степень гнездования, “разбросав” выборку по большому числу более разобщенных территориальных единиц. Но все это ведет к увеличению затрат и может повлечь за собой смещения, не связанные с выборкой, так как затрудняется контроль качества и наблюдение за ходом возросших масштабов работ. Поэтому необходим баланс, который позволил бы минимизировать ошибки при данных затратах.

Планирование выборки и процедуры оценки могут особенно усложниться при непрерывных обследованиях рабочей силы, когда следует учитывать пересечения и ротацию выборки, одновременную представительность в пространстве и во времени, необходимость периодического пересмотра выборки и много других факторов (см. гл. 10 и 11). Очевидно, что многие из этих решений

невозможно принять на научной основе, если отсутствует подробная информация об ошибках выборки в зависимости от ее объема и плана.

Оценка

Число оценок, полученных в рамках типичного обследования рабочей силы, может быть весьма большим. При непрерывном обследовании некоторые показатели ошибки выборки могут фактически быть довольно стабильными от раунда к раунду, так что нет необходимости повторять расчеты ошибок в каждом раунде. Тем не менее методы исчисления ошибок выборки должны быть достаточно простыми и выражаться в легких расчетных формулах, что позволит экономично провести подробные расчеты, требующиеся для многочисленных оценок. В то же время методы расчета должны быть достаточно универсальными, чтобы их можно было применить для сложных вариантов, которые часто встречаются в планах выборки для крупномасштабных обследований.

Обзор практических методов расчета ошибок выборки содержится в работе (Kalton, 1977). Один из наиболее часто применяемых методов базируется на следующем принципе. При некоторых допущениях, обычно не очень ограничительных в практических условиях, ошибки выборки для различных показателей — долей, средних, коэффициентов и их разностей, для выборки в целом и отдельных подклассов можно получить на основе итоговых данных на уровне первичных единиц отбора. План выборки и отбор внутри отдельных первичных единиц могут быть сложными и отличаться друг от друга, но это не влияет на оценку дисперсии. При таком подходе *фактически* учитывают элементы дисперсии на всех ступенях многоступенчатого плана, включая вторую и последующую ступени, хотя в вычислительной формуле никаких явных ссылок на другие ступени, помимо первой, не содержится. Суть дела в том, что дисперсия, вносимая на последующих ступенях, отражается в наблюдаемой вариации выборочных результатов на первой ступени отбора. При этом требуются следующие основные допущения: а) из каждого слоя извлекают две или более первичные единицы отбора, б) этот отбор независим друг от друга, случаен и производится с возвращением. Эти условия при отборе на практике редко удовлетворяются полностью, но многие ситуации приближаются к ним. Исходя из названных допущений оценки дисперсии для линейных функций характеристик, например итоговых оценок для обследования, можно получить с помощью аналитических формул. Для нели-

нейных функций (подобно оценкам коэффициентов во многих обследованиях) обычно прибегают к классу методов, именуемых “методами разложения в ряд Тейлора”. Разработаны и достаточно широко применяются численные методы и компьютерные программы для этого подхода (см., например, Woodruff and Causey, 1976; Verma, 1982; Verma and Pearce, 1987; Schnell et al., 1988). Существуют также другие методы исчисления ошибок выборки для комплексных данных на основе сложных выборок, например, “сбалансированные повторные опыты” и “jack-knife повторные опыты”; численные примеры и обзор методов см. (Kish and Frankel, 1974). Общий анализ оценки дисперсии содержится у (Wolter, 1985).

Характер изменчивости

По ряду причин целесообразно изучать характер изменчивости результатов ошибок выборки в разрезе показателей обследования, классов выборки, отдельных обследований и раундов и соотносить эту изменчивость со структурой выборки.

1) *Экстраполяция полученных результатов.* Оценки, представляющие интерес в рамках крупномасштабного, многоцелевого и многораундового обследования, как правило, слишком многочисленны, чтобы можно было исчислить ошибки выборки для каждой из них. Необходимо иметь метод экстраполяции ошибок для некоторых показателей и категорий выборки на другие показатели и категории, для которых расчеты не производились. Это предполагает изучение характера изменчивости ошибок выборки в разрезе показателей и категорий. Особенно целесообразно это при непрерывных обследованиях, когда некоторые ошибки выборки обнаруживают стабильность от раунда к раунду, так что на основе ранее полученной картины можно предсказать ошибки выборки в последующих раундах.

2) *Обобщение при представлении результатов.* Даже если бы были исключены ошибки выборки для всех опубликованных оценок показателей, публиковать их было бы нецелесообразно. Обычно информация о выявленной картине ошибок требует обобщения.

3) *Сглаживание полученных результатов.* Ошибки выборки, исчисленные на основе выборочных данных, сами подвержены значительным колебаниям, особенно для выборок с небольшим количеством первичных элементов. Поэтому часто предпочтительнее и правильнее полагаться на должным образом исчисленные средние по нескольким расчетам, а не на точность отдельных расчетов.

4) *Планирование и уточнение выборки.* Расчеты ошибок выборки призваны показать не только степень точности полученных оценок, но и пути пересмотра плана выборки в будущем. Для указанной цели необходимо соотнести характер ошибок со структурой выборки, например, гнездованием, стратификацией (расслоением) и взвешиванием. Взаимосвязь между ошибками и структурой выборки важна также с позиций экстраполяции, обобщения и сглаживания исчисленных результатов.

Распространение результатов

Чтобы достичь указанных целей, необходимо найти такие показатели изменчивости отбора, которые позволили бы распространить полученные результаты с одной категории выборки по другую, с одного показателя на другой, с одного обследования и раунда на другие.

На основе картины ошибок выборки, исчисленных для различных оценок при разных объемах выборки на базе либо данного обследования, либо прошлых раундов или аналогичных обследований, можно построить примерные модели или графики, позволяющие исчислить или установить примерные значения ошибок выборки для других показателей. (Подробные иллюстрации см. в работах Gonzalez et al., 1975; Verma, 1982).

Величина ошибки выборки для данного показателя зависит от ряда факторов: природы показателя, единицы его измерения и типа оценки; объема и плана выборки; характера и распределения представляющих интерес подгрупп в выборочных гнездах; в многораундовых обследованиях — также от схемы пересечения и ротации выборок. Для контроля и уменьшения последствий влияния этих факторов используют различные меры стандартных ошибок, что также позволяет распространить эти измерители на другие случаи. Ниже приведены некоторые распространенные меры, полученные на основе стандартной ошибки (которая представляет собой квадратный корень из дисперсии):

1) *Коэффициент вариации.* Он представляет собой стандартную ошибку, деленную на среднее значение показателя, и может быть стабильным в ходе повторяющихся обследований одинаковой схемы, объема и содержания. Разделив стандартную ошибку на среднее значение показателя, исключают влияние единицы измерения, масштаба и величины оценки. В той мере, в какой коэффициенты вариации, исчисленные для одного набора данных, удается использовать для другого множества данных, потребность в новых расчетах

сокращается (более подробные и практические иллюстрации содержатся в работе Zarkovich, 1979.)

2) *Фактор плана*. Это одна из самых полезных мер в нашем контексте, и она характеризует общее влияние структуры выборки на величину ошибки отбора. Определяется как соотношение стандартной ошибки для реализованного плана выборки (se) и стандартной ошибки простой случайной выборки того же объема (sr):

$$deft = se/sr$$

(Часто используют равноценную меру “эффект плана”, или $deff$, т.е. попросту квадрат $deft$, иначе говоря отношение фактической дисперсии и дисперсии простой случайной выборки). $Deft$ — это более универсальная мера, нежели se (стандартная ошибка), ибо она не зависит от факторов, влияющих на se и sr , например, от единицы измерения, величины оценки, степени ее изменчивости в совокупности и, прежде всего, — от объема выборки. Но $deft$ зависит от других факторов, таких, как природа оцениваемого показателя, план выборки, тип и размер элементов выборки. $Deft$ — обобщающая мера последствий отклонения фактического плана выборки от простой случайной выборки учтенных единиц. Это всеобъемлющий фактор, с помощью которого пытаются обобщить влияние различных тонкостей плана, особенно гнездования, расслоения и взвешивания. Он способен даже уловить влияние оценки по уравнению регрессии, двойного отбора, ротации и пересечения выборок от раунда к раунду и т.д. По указанным причинам многие специалисты обычно включают отношение se/sr в выдаваемые расчеты дисперсии.

$Deft$ можно также рассматривать в качестве показателя того, как гнездование и другие аспекты плана (которые позволяют снизить затраты и улучшить контроль за другими ошибками, не связанными с выборкой) фактически ведут к уменьшению объема выборки. Если обозначить через n фактический объем выборки, то при простом случайном отборе ту же ошибку выборки дал бы ее меньший объем (n'), где

$$n' = n/deft^2$$

Несколько эмпирических иллюстраций и их обсуждение содержатся у (Kalton and Blunden, 1973; Kish et al., 1976; Verma et al., 1980).

3) *Коэффициент однородности (roh)*. Для данного показателя и при данном числе и типе гнезд, а также данной процедуре отбора подвыборки величина $deft$ по мере увеличения размера гнезд, как правило, возрастает. Чтобы учесть это влияние, Киш (1965) пред-

ложил сводную меру roh (коэффициент однородности), определяемый как

$$deft^2 = 1 + (b - 1)roh,$$

где b — “средний объем гнезда” (т.е. количество учтенных единиц, отобранных в расчете на одну первичную единицу выборки). Модель основана на понятии внутригрупповой корреляции, т.е. корреляции между элементами гнезда. Приведенное выше уравнение предложено для самовзвешивающихся выборок при отсутствии чрезвычайных колебаний в объеме гнезд, хотя ее можно адаптировать применительно к более общим условиям. Коэффициент однородности — это синтетическая мера, предложенная с целью измерить среднюю степень однородности значений определенного показателя единиц первичного отбора в сопоставлении с общей изменчивостью этого показателя. В принципе, roh более универсален, нежели $deft$, так как в нем элиминировано влияние различий объема выборки в расчете на гнездо. Однако на практике рассчитанная величина roh , как правило, нестабильна и изменчива.

4) *Корреляция между раундами.* При непрерывных обследованиях может также оказаться весьма полезным накопление информации о корреляции фиксированного набора единиц (или единиц из каким-то образом связанных выборок) между раундами, а также о других аспектах ошибок отбора в оценках изменений от раунда к раунду. Такая информация существенна для улучшения процедур оценки, схем ротации и различных аспектов плана выборки при продолжающихся обследованиях. (См., например, обсуждение составных оценок при пересекающихся выборках в парагр. 6 гл. 11.)

Пример.

Приводимый ниже пример из обследования рабочей силы в Австралии показывает простой и компактный метод представления информации об ошибках отбора, очевидно, основанный на результатах продолжающегося обследования (ILO, 1986). В табл. 21 представлены величины стандартной ошибки в виде доли в величине оценки (числе лиц в данной категории) как функция последней.

Таблица 21. Относительная стандартная ошибка как функция величины оценки

Величина оценки (количество лиц в категории)	Относительная ошибка (% оценки)	Величина оценки (количество лиц в категории)	Относительная ошибка (% оценки)
4500	21,6	200 000	2,6
5000	20,0	300 000	2,0
6000	18,0	500 000	1,4
10 000	14,0	1 000 000	0,9
20 000	10,0	2 000 000	0,6
50 000	5,8	5 000 000	0,3
100 000	3,9		

Приведенную таблицу можно использовать для определения доверительных интервалов при оценке величины (числа лиц) различных категорий — занятых, безработных, лиц, работающих на собственных предприятиях, и т.п. Для отдельных показателей приблизительные значения относительной стандартной ошибки получают путем умножения приведенных выше цифр на множитель в зависимости от показателя. Применительно к обследованию рабочей силы в Австралии эти множители определены в следующих размерах:

Показатель	Множитель (на который умножают приведенные стандартные ошибки)
Общее количество отработанных часов	1,2
Среднее количество отработанных часов	0,5
Средняя продолжительность безработицы	1,5
Медианная продолжительность безработицы	2,0

5. ОШИБКИ В ОТВЕТАХ

Характер и источники ошибок в ответах

Ошибки в ответах появляются на этапе сбора данных в результате неточного определения и формулирования вопросов, а затем в ходе подачи вопросов, получения и регистрации ответов. Ошибки, возникающие на последующих этапах обработки ответов, обсуждаются в следующем параграфе. На практике из-за того, что ответы можно изучить и проанализировать лишь после некоторой обработки (кодирования, ввода данных и т.п.), ряд ошибок, появляющихся на этапах обработки, часто путают с ошибками в ответах. Точно так же некоторые элементы ошибок в ответах смешивают с обычными оценками ошибок отбора, так как последние базируются на величинах, которые можно получить лишь после наблюдения. Это иллюстрирует положение, высказанное в парагр. 1, а именно, что компоненты ошибок, измеренные различными методами, нельзя ни вычитать, ни складывать, чтобы получить общую величину совокупной ошибки. Это положение следует иметь в виду при последующем обсуждении проблемы.

Проблема ошибок в ответах многогранна, так как источники, характер и форма ошибок многообразны и меняются в зависимости от содержания и условий проведения обследований.

В каком-то самом глубоком смысле ошибки в ответах возникают в связи с недостатками *содержательного аспекта* обследования:

понятия и подлежащие учету характеристики следует преобразовать в набор вопросов и категорий, которые должны учитывать все разнообразие встречающихся условий, включая многие пограничные ситуации, которые четко не укладываются в одну концептуальную категорию. Различные концептуальные проблемы в связи с качеством ответов при обследованиях рабочей силы рассмотрены в части 1 настоящего руководства. Так, в гл. 2 отмечалось, что субъективное толкование респондентами и интервьюерами понятий “работа” и “экономическая активность” едва ли будет таким всеохватывающим, как это вытекает из определений обследования. Было показано, как уточнение в ходе опроса с помощью подробного перечня видов деятельности способно повлиять на полную выявления экономической активности. В гл. 2 обсуждались также трудности при определении статуса активности в пограничных случаях, т.е. в ситуациях, когда небольшое изменение одного или нескольких аспектов понятия рабочей силы приводит к изменению статуса респондента. В гл. 3 и 4 рассматривается, соответственно, как путем выяснения активности день за днем можно повысить точность данных об активности в данный момент, а путем помесечного анализа — улучшить данные об обычной активности.

Ошибки в ответах возникают также при *реализации плана* обследования в результате ошибок со стороны респондента, интервьюера или того и другого. Одна из очевидных причин — неспособность респондента дать требуемую информацию: респондент, возможно, просто никогда не знал некоторые факты, а если теоретически они ему известны, он может быть не в состоянии вспомнить их в момент опроса, особенно если они второстепенны и не запечатлелись в памяти. Респонденты могут также не желать сообщать некоторую информацию (например, о деятельности, которая не вполне легальна или осуществляется нелегальными методами), либо сообщать верную информацию по той или иной причине. Еще один источник неопределенности возникает в силу того, что во многих обследованиях рабочей силы вынуждены довольствоваться определенной долей ответов, данных за другое лицо. Даже когда такие ответы не очень влияют на чистые одномерные распределения, установлено, что они могут привести к более существенным валовым расхождениям и серьезно исказить многоаспектные таблицы и анализ данных (см., например, Martin and Butcher, 1982). Точно так же интервьюеры могут явиться причиной ошибок, если они не понимают или неправильно применяют понятия и процедуры, руководствуются субъективными предпочтениями, плохо регистрируют ответы, а иногда даже поступают бесчестно.

Учитывая разнообразие и сложность ошибок в ответах, их оценка и контроль требуют концентрации и экономии усилий. Цель контроля должна состоять в том, чтобы: а) выявить наиболее существенные и распространенные виды ошибок, встречающихся в той или иной ситуации, б) подсказать, насколько это возможно, конкретные причины ошибок и источники их возникновения, с тем чтобы можно было сосредоточить усилия и средства на наиболее эффективных мерах по улучшению качества ответов. Отсюда следует, что оценки и анализ ошибок в ответах следует проводить на уровне отдельных респондентов и интервьюеров, а также на агрегированном уровне с точки зрения их влияния на показатели обследования. Изучение на индивидуальном уровне поможет выявить конкретные причины и обстоятельства возникновения ошибок, а анализ на более агрегированных уровнях — значимость и влияние ошибок на результаты.

Анализ ошибок на индивидуальном уровне: ошибка в ответах на этом уровне означает расхождение между полученным ответом и подлинным значением показателя. Это может принимать различные формы: отсутствие ответа на пункт, неправильная информация, неполные данные, неправильное отнесение событий по времени и т.д.

Эффективный путь выявления полезной информации о характере, источнике и частоте ошибок в ответах (и при обработке), влияющих на конечные результаты обследования, заключается в *анализе индивидуальных записей*, где могли возникнуть ошибки. Эти случаи можно выявить на основе заранее установленных моментов, например, при наличии необычных или маловероятных кодов и сочетаний характеристик. Затем можно проверить анкеты или другие документы, с тем чтобы установить, возникла ли ошибка при обработке (кодирование, ввод данных и т.д.) или же на этапе сбора данных. Независимо от причины ошибки, задача состоит в анализе отдельных случаев с целью выявить, где и как могла появиться ошибка. Вообще говоря, при таком анализе основная цель состоит не в том, чтобы получить количественные меры качества агрегированных данных, каковыми являются дисперсия или смещение.

Смещение оценок

По характеру влияния на результаты обследования на агрегированном уровне полезно различать смещение и изменчивость ошибок в ответах. Дело в том, что оба эти элемента ошибок принципиально различаются по методам их оценки и контроля. Как указано в па-

разр. 1, смещения — это систематические ошибки, которые являются следствием общих условий проведения обследования. Их, как правило, можно установить путем сопоставления с информацией и источниками, внешними по отношению к самому обследованию. Так, при непрерывном обследовании рабочей силы, где применяют более или менее стандартизованные схемы и процедуры из одного раунда в другой, на результаты всех раундов могут повлиять в сущности одни и те же смещения в ответах. Следовательно, такие смещения могут оказаться незамеченными, если результаты не сопоставить с внешними данными.

Первый шаг при выявлении смещения заключается в логическом и содержательном анализе непротиворечивости и взаимосвязей данных при их сопоставлении с внешними стандартами и ранее имевшейся информацией данному вопросу. Кроме того, для установления размера смещения требуется сопоставить результаты обследования с более точной информацией. Последняя может содержаться в каком-либо независимом надежном внешнем источнике или может быть получена на основе специальных операций по сбору данных. В последнем случае имеется несколько возможностей, рассмотренных ниже.

Повторный опрос на основе схожих, но улучшенных процедур

Идея состоит в получении в сущности той же информации, что и при первоначальном обследовании, и на базе схожих, но усовершенствованных процедур. “Схожесть” процедур заключается в том, что источник информации будет того же типа, что и при первоначальном обследовании; либо будет опрошен тот же респондент, либо, когда это допускается, другой член того же домашнего хозяйства. Опрос базируется на той же анкете, но она может быть дополнена некоторыми вопросами и уточнениями для ряда пунктов. Процедуры можно улучшить, к примеру, путем использования более опытных и лучше обученных интервьюеров, более тщательного наблюдения и контроля, дополнительных вопросов и уточнений, всесторонней проверки пунктов, где возможны или были выявлены ошибки.

Базовая модель (метод 1) заключается в проведении двух опросов в подвыборке после проведения первоначального опроса. Они включают: а) повторный опрос другим интервьюером как повторение первоначального опроса, после чего следует б) третий “согласительный” опрос, призванный выявить и снять расхождения между пер-

выми двумя опросами, получить правильные ответы и установить причины сделанных ранее ошибок. Третий опрос необходим лишь тогда, когда два первых не согласуются между собой и явно требуют улучшенных и более интенсивных подходов. Цель повторных опросов — получить информацию о дисперсии ответов (см. ниже); они помогают также установить случаи расхождений для последующей проверки и тем самым сократить число более дорогостоящих “согласительных” опросов. Задача последних — установить подлинные величины показателей для измерения смещения в ответах и причины расхождений на микроуровне.

Потенциальное преимущество охарактеризованной модели — это возможность изучить причины ошибок в ответах на индивидуальном уровне и получить меры дисперсии и смещения. Возможные проблемы этого метода: сравнительно высокие затраты в связи с двумя дополнительными опросами другими интервьюерами; трудности в подыскании квалифицированных работников для этой цели и обеспечении того, чтобы опросы проходили независимо от первоначального и с соблюдением, в сущности, условий первого опроса. Этот метод не способен также уловить случаи ошибок, когда они присущи и первоначальному и повторному опросам, так как при этом не возникают расхождения, требующие дополнительного изучения.

Возможны варианты указанной базовой модели. Один из них (метод 2) также заключается в проведении независимого повторного опроса, но с совмещением его с “согласительным” опросом в рамках одной операции для экономии затрат (и времени). Этот метод используют, например, организаторы текущего обследования населения в США (United States, 1978). “Опасность” при таком подходе состоит в том, что факт последующего “согласительного” опроса может повлиять на поведение интервьюера при повторном опросе; он может также попытаться “подправить” повторный опрос на базе последующего согласительного опроса. Чтобы уменьшить остроту этих проблем, можно предпринять два шага (ibid, p. 31): а) те, кто проводит повторный опрос, воздерживаются от ознакомления с ответами первоначального опроса до завершения повторного опроса и их инструктируют, указывая на недопустимость вносить какие-либо изменения в анкеты повторного опроса на основе “согласительной” операции; б) определенная доля повторных опросов последующему согласованию не подвергается. По идее, проводящий повторный опрос не знает, потребует ли согласительный опрос для данной единицы, пока повторный опрос не завершен. С течением времени доля согласительных операций в рамках текущего обследо-

вания населения была увеличена до 80%, так как они обеспечивают ценную информацию о смещении и источниках ошибок. В канадском обследовании рабочей силы следуют аналогичной схеме, причем согласование проводят по отобранной случайным способом половине повторного опрошенных единиц (тем же интервьюером), а другая половина согласительному опросу не подвергается (Canada, 1976, р. 84). Накопление информации из обеих частей выборки позволяет изучить общие аспекты ошибок в ответах. Простую дисперсию ответов можно оценить на базе несогласованной части анкет при предположении, что опрос был повторен при тех же условиях, а смещение — на основе согласованных анкет при предположении, что в результате согласительного опроса получены “истинные” ответы. Можно также изучить общие аспекты работы интервьюеров, проанализировав результаты с учетом характеристик работников.

При другом варианте (метод 3) попытки провести независимые повторные опросы не предпринимаются. Проводят лишь два опроса. В ходе второго опроса интервьюер имеет перед глазами ответы, полученные при первоначальном опросе. Если ответы в двух опросах расходятся, интервьюер немедленно пытается, как можно тактичнее, выяснить причину расхождений и записать правильный ответ. Часто при втором опросе прибегают к улучшенным процедурам (например, дополнительным уточнениям и вопросам, с тем чтобы повысить эффективность выявления ответов). Недостатки этого метода таковы: а) неспособность измерить *дисперсию* ответов; б) специально обученные интервьюеры, способные примирить расхождение и найти правильные ответы, требуются для всего повторного опроса, а не только для согласительных вариантов, как при других методах. Преимуществами его являются: а) экономия средств, так как проводят два, а не три опроса; б) более четкая ориентация на выявление смещения и конкретных причин ошибок; в) меньшая вероятность того, что первоначальному и повторному опросам присущи одни и те же ошибки, остающиеся незамеченными и неучтенными. Примером использования такого метода может служить повторное обследование, проводимое вслед за раундом обследования рабочей силы в Швеции (Sweden, 1980b).

Еще один метод (4) заключается в проведении независимых повторных опросов, но без согласительной части. Независимый повторный опрос можно проводить: а) при, в сущности, тех же условиях, что и первоначальный опрос, так чтобы он примерно повторял его, либо б) независимо и на основе улучшенных методов. В первом случае схема позволит измерить простую дисперсию ответов, но не смещения; расхождения в ответах не укажут на возмож-

ный источник ошибок. Как упоминалось выше, эта альтернатива используется в *части* (несогласованных) выборочных повторных опросов, например, в текущем обследовании населения в США и обследовании рабочей силы в Канаде. Второй подход ориентирован на измерение смещения постольку, поскольку повторный опрос дает правильные ответы. Независимость повторных опросов это не строгая теоретическая предпосылка для указанной цели, но она считается желательной при предположении, что при опросе тем же интервьюером вероятны систематические смещения при выявлении случаев ошибок в ответах.

Описанные выше варианты отличаются по своим предположениям относительно характера процесса опроса и повторного опроса. В отличие от метода 3, методы 2 и 4 основаны на предположении, что ошибки можно выявить более эффективно на основе независимого повторного опроса. Но что касается согласования и установления истинных ответов, методы 2 и 3 исходят из того, что это можно сделать эффективнее путем проверки выявленных случаев расхождений, тогда как метод 4 базируется на предположении, что истинные ответы можно получить с помощью независимых повторных опросов на основе улучшенных процедур и что такой подход менее чреват пристрастным отношением интервьюеров по сравнению с независимым согласительным опросом. Имеющихся фактов недостаточно, чтобы определенно рекомендовать какой-либо один из перечисленных методов. Любопытно отметить, что в рамках обследования рабочей силы в Швеции прибегают к методам 3 и 4 с весьма схожими результатами (Sweden, 1980b).

Сопоставление исследований, основанных на различных методологиях

Иногда одну и ту же информацию можно собирать с использованием различных подходов, вопросов, отчетных периодов, типов респондентов, источников информации, методов сбора данных и т.п. Если результаты при двух различных подходах существенно разнятся, возникает сомнение в обоснованности одного из них или обоих. Если априори известно, что один подход лучше, тогда такое сопоставление способно по меньшей мере указать на смещение в худшем (и обычно менее дорогостоящем) подходе. Весьма полезно заранее знать *направление* ожидаемого смещения. Так, если по результатам обследования ожидается заниженная оценка определенного показателя из-за пропусков, тогда процедуру (скажем, на основе вспоминания день за днем отчетной недели), которая обес-

печивает большую величину оценки, можно считать менее смещенной, нежели другую процедуру (скажем, основанную на воспоминаниях целой недели), которая дает меньшую величину. Так, Ниemi (1983 г.) проанализировал смещения в полученных данных о рабочем времени в ходе обследования рабочей силы в Финляндии, сопоставив их с результатами более глубокого изучения подвыборки на основе баланса использования времени и подробных дневников респондентов. Эти дневники, хотя и они не свободны от ошибок, можно считать более точными. Поэтому такое сопоставление позволяет судить о смещении оценок, присущем обследованиям рабочей силы. Точно так же Меллоу и Сайдер (1983 г.) проанализировали ответы на вопросы об отрасли, занятии, профсоюзном членстве, рабочем времени и заработной плате путем сопоставления информации, полученной от рабочих, с данными по тем же рабочим, представленными работодателями в специальном приложении к текущему обследованию населения в США.

Дисперсия ответов

Дисперсия ответов обусловлена преходящими или случайными факторами, влияющими на отдельные ответы (в отличие от смещений, которые более или менее систематически относятся ко всем респондентам). Если можно было бы повторить обследование на основе тех же процедур и существенных условий его проведения, результаты при каждом повторении были бы разными даже при той же выборке. Дисперсия ответов и является мерой этой изменчивости.

Различают две компоненты ошибки изменчивости: простую дисперсию ошибок и коррелированную дисперсию ошибок. Такое различие полезно и важно, так как компоненты неодинаковы по своей природе и методам оценки. Несколько упрощенно эти понятия можно охарактеризовать следующим образом.

Любому интервьюеру (а также бригадиру, кодировщику или другому работнику) могут быть присущи пристрастия, которые влияют на всех респондентов, опрос которых включен в круг его обязанностей. Несущественно, являются ли эти пристрастия следствием ошибок или невежества со стороны интервьюеров или респондентов. Некоторые смещения могут быть общими для результатов работы всех интервьюеров, отражая общие условия обследования и тип задействованных работников. Это соответствует смещению в ответах, охарактеризованному выше. Но поскольку отдельные работники оказывают *различное* среднее влияние на результа-

ты своей работы, в пределах этих результатов ошибки коррелируют. Коррелированная дисперсия — это мера изменчивости, связанная с таким влиянием. Помимо коррелированных ошибок, обусловленных различным средним влиянием отдельных работников, имеются также случайные факторы, которые приносят случайные ошибки в полученные ответы независимо от конкретного интервьюера, проводящего опрос. Это вторая компонента отражается в простой дисперсии ответов (см. Hansen et al., 1961; Kish, 1962).

Простая дисперсия ответов

Простая дисперсия ответов — это показатель внутренней нестабильности (неповторяемости) полученных ответов. Он может быть полезен при оценке доходчивости отдельных пунктов анкеты. Если простая дисперсия ответов по какому-то пункту значительна, это может означать, что полученная информация относительно ненадежна, возможно, из-за проблем в содержании и формулировках анкеты.

Измерение простой дисперсии ответов предполагает независимый *повторный* сбор данных при тех же общих условиях обследования. В рамках одного обследования невозможно провести различие между варьированием истинных значений индивидуальных показателей единиц совокупности (что порождает ошибки выборки) и их дополнительной изменчивостью, связанной с влиянием случайных факторов на отдельные ответы (что ведет к простой дисперсии ответов). Обычная процедура исчисления ошибки выборки на деле автоматически учитывает и полное влияние простой дисперсии ответов. Обе эти компоненты дисперсии одинаковым образом уменьшаются по мере увеличения объема выборки.

Отдельная оценка простой дисперсии ответов требует повторных опросов. Их следует проводить независимо от первоначального обследования, но при тех же условиях и с использованием тех же методов. На практике точно выполнить оба эти условия трудно.

Можно построить различные обобщенные меры простой дисперсии ответов. Одна из них — “коэффициент несовместимости”, определяемый как отношение простой дисперсии ответов и общей дисперсии — простой и выборочной; он характеризует увеличение дисперсии элементов в связи со случайной нестабильностью полученных ответов. Эта мера наиболее полезна применительно к таблице 22 “опрос — повторный опрос” для дихотомического показателя. В табл. 22 показано, как исчисляют коэффициент несовместимости и некоторые другие простые меры.

Обобщенные меры, подобно коэффициенту несовместимости, не столь полезны применительно к многопризначным показателям. В этом случае целесообразнее проанализировать всю комбинационную таблицу ответов при опросе и повторном опросе с целью выявить признаки (категории), которые особенно подвержены изменчивости в ответах. Примерами таких комбинационных таблиц на основе данных обследований рабочей силы служат таблицы опроса — повторного опроса, касающегося статуса, основного вида активности за отчетную неделю, отрасли, занятия и группировки по отработанным часам.

Т а б л и ц а 22. Исчисление коэффициента несовместимости на основе повторного обследования

Отношение к рабочей силе в первоначальном обследовании	Отношение при повторном обследовании		
	в составе рабочей силы	вне рабочей силы	итого
В составе рабочей силы	a	b	a + b
Вне рабочей силы	c	d	c + d
Итого	a + c	b + d	n = a + b + c + d
Отношение разности (чистое) $e = (a + c) / n - (a + b) / n = (c - b) / n$			
Отношение разности (валовое) $g = (c + b) / n$			
Средняя доля (единиц), обладающих определенной характеристикой $p = [a + (c + b) / Z] / n$			
Коэффициент несовместимости $I = g / Zp (1 - p)$			
<i>Примечание:</i> a — d это полученные в результате наблюдения частоты в клетках комбинационной таблицы "опрос-повторный опрос".			

Коррелированная дисперсия ответов

Как уже отмечалось, коррелированная дисперсия ответов (или "дисперсия интервьюера") вызвана различным влиянием интервьюеров на результаты их работы. Она указывает на отсутствие единообразия и стандартизации в их работе. Высокие значения этого показателя свидетельствуют о необходимости лучшего обучения и контроля за работой интервьюеров. Его величина находится в обратном отношении к числу интервьюеров, занятых в обследовании, подобно тому как величина ошибки отбора зависит от числа гнезд в выборке.

Важно поэтому сознавать, что точно так же, как ошибки отбора могут быть довольно значительными для оценок, основанных на небольшом числе первичных единиц отбора, влияние дисперсии ответов может оказаться существенным при оценках, полученных с небольшим числом интервьюеров. Это важное соображение, когда решается вопрос, должным ли образом организовано обследование для получения оценок на уровне регионов или провинций. Если интервьюеры не прошли одинаково хорошую подготовку и их работа не контролируется, то связанные с ними ошибки могут превзойти по значимости все другие источники ошибок в региональных оценках, особенно, когда каждая из них основывается на работе немногих интервьюеров, даже если эта проблема на общенациональном уровне не существует (United Nations, 1982, p. 156—157). Поэтому теория гласит, что работу по опросу следует распределить между возможно большим числом интервьюеров. Но на практике следует сопоставить более высокую точность с затратами на набор, обучение, контроль за работой и разъезды интервьюеров. Кроме того, возможности организации провести хорошее обучение и обеспечить должный контроль могут оказаться ограниченными. Слишком большое количество работников может привести к ухудшению качества, а следовательно, к значительным смещениям в оценках, что способно перечеркнуть любой выигрыш в эффективности, связанный с уменьшением нагрузки на интервьюера. Однако следует избегать и крайних случаев — использования небольшого числа интервьюеров в крупномасштабных обследованиях. Заметим, что эти соображения весьма схожи с теми, которые связаны с определением гнездования выборки при многоступенчатом отборе.

Измерение коррелированной дисперсии ответов предполагает сопоставления между должным образом спроектированными *повторными* обследованиями или взаимопроникающими выборками, т.е. между разными частями выборки, каждая из которых относится к работе одного интервьюера и дает по-своему обоснованную оценку. Главное требование относится к степени *рандомизации* (случайности) при распределении элементов выборки между отдельными интервьюерами. В каком-то смысле интервьюеры приносят свое гнездование в обследуемую выборку из-за своих индивидуальных пристрастий. В той мере, в какой это гнездование совпадает с географическим гнездованием самой выборки (например, если используют постоянные счетчиков — по одному на каждое гнездо выборки), обычная оценка ошибки отбора автоматически вберет в себя коррелированные ошибки ответов. С другой стороны, если распределение опросов в пределах гнезд выборки и между ними носит

случайный характер (рандомизировано), влияние фактора коррелированных ошибок не найдет отражения в оценке ошибки отбора, а может быть исчислено отдельно с помощью аналогичной процедуры, причем нагрузки интервьюеров рассматриваются в качестве показателя “первичного отбора”.

Такой же подход фактически применим и к анализу коррелированного влияния бригадиров, контролеров, кодировщиков и т.д. При оценке отдельной компоненты коррелированной ошибки ее величина будет завышена в той мере, в какой другие ее источники не носят случайного характера, а следовательно, их путают с измеряемой компонентой. Вполне завершенная схема организации могла бы предполагать простой случайный отбор равного числа опросов для всех интервьюеров и распределение работ между бригадами, контролерами и кодировщиками в виде ортогональной (прямоугольной) системы, с тем чтобы можно было отдельно исчислить все компоненты. Однако случайное распределение нагрузки ведет к увеличению затрат и за определенными пределами часто становится нецелесообразным. К счастью, полезные оценки коррелированных ошибок обычно можно получить, даже если рандомизация загрузки работников не является законченной (Kish, 1962).

6. ОШИБКИ ПРИ ОБРАБОТКЕ

Ошибки возникают также на различных этапах обработки данных: кодирования, ввода данных, редактирования, добавления отсутствующих данных, табулирования и т.д. Настоящий раздел посвящен оценке некоторых мер ошибки, прежде всего связанных с кодированием.

Оценка ошибок кодирования

Кодирование — это операция, в ходе которой данные из анкет преобразуют в форму, пригодную для ввода в компьютер. Работа по кодированию существенно различается по своей сложности — от преобразования сложных словесных описаний в числовые коды, требующего значительного опыта и знаний, до простого переписывания имеющихся числовых данных в предназначенные для них клетки.

Ошибки кодирования можно классифицировать таким же образом, как и ошибки в ответах, рассмотренные в предыдущем разделе. Систематические ошибки, присущие работе всех кодировщиков, — это *смещения*. Они могут быть обусловлены, например, определенным изъяном в системе или процедуре кодирования. Каждый кодировщик также оказывает свое особое влияние на результаты работы.

Это различное влияние отдельных кодировщиков порождает *коррелированную дисперсию кодирования* (или проще “дисперсию кодирования”). Имеются также случайные ошибки кодирования ответов, не связанные с нагрузкой данного кодировщика: *простая дисперсия кодирования*.

На масштабы ошибок кодирования влияют многообразные факторы. Помимо самой природы кодируемого пункта, существенным фактором является построение анкеты, например, наличие открытых, а не закрытых вопросов, а также способ кодирования в анкете. Еще один важный фактор — организация и контроль операции кодирования, степень централизации или децентрализации. Централизованное и более тщательно контролируемое кодирование может иметь ряд преимуществ: меньшую изменчивость; возможность создания более полной основы кодирования благодаря наличию более полного набора ответов; лучшие условия для повторения и проверки работы. В то же время децентрализация кодирования на местах имеет преимущество в экономии средств и скорости. Точно так же различны характер и масштабы ошибок при использовании обученных кодировщиков, кодировании интервьюерами, самокодировании респондентами или теми, кто дает за них сведения. Меняются они и в зависимости от демографических характеристик респондентов и т.п. По вопросу об ошибках кодирования существует обширная литература, но обычно применительно к конкретным условиям развитых в статистическом отношении стран. В частности, можно порекомендовать два исследования: Jabine and Terping, 1973; Lyberg, 1982. В обоих внимание сосредоточено на кодировании данных об отрасли и занятии, которые, пожалуй, представляют собой наиболее (а часто единственно) важные и сложные пункты, требующие кодирования при обследованиях рабочей силы.

Качество кодирования можно оценить на основе: а) информации о контроле качества, полученной в рамках обычной операции кодирования; б) анализа результатов случайного распределения нагрузки между кодировщиками; в) сопоставления с результатами повторного кодирования, обычно в виде отдельной операции и на выборочной основе.

Цель случайного распределения (б) — измерить коррелированную дисперсию. Процедура аналогична той, которая охарактеризована выше применительно к дисперсии опроса интервьюера. Главное различие в том, что случайно распределить нагрузку между кодировщиками значительно легче, чем нагрузку интервьюеров, и это не связано с серьезными организационными и финансовыми трудностями.

Цель повторного кодирования (в) — выявить ошибки и их причины на индивидуальном уровне, установить смещение и ошибки изменчивости. Как и при анализе ошибок ответа, повторное кодирование может быть независимым и зависимым, и может проводиться, в сущности, при тех же условиях, что и первоначальное кодирование, или на основе усовершенствованных процедур. Рассмотренные выше различные альтернативы применимы и здесь.

Что же касается операций контроля качества (а), можно иметь подробную информацию о разных мерах, например, о коэффициентах входных ошибок, общих коэффициентах проверки, доле бракованных партий и т.д. Их можно сгруппировать по отдельным кодировщикам и сопоставлять во времени при продолжающихся обследованиях. Особое внимание следует уделить пунктам, которые считаются сложными, например, занятию и отрасли. Упомянутые показатели рассмотрены ниже.

Простейший способ организации кодирования и контроля качества состоит в следующем. Подлежащие кодированию документы делаются на пакеты (партии). С учетом требований контроля качества в каждой партии проверяют определенное число записей. Если число ошибок в партии не превышает некий заранее установленный уровень, партию принимают. В противном случае ее “бракуют”, т.е. подвергают 100%-ной проверке.

1) Коэффициент входной ошибки — это доля случаев, содержащих ошибки после первоначальной операции кодирования при отсутствии мер контроля (верификации). В каждой партии, проверенной на выборочной основе, число общих ошибок можно исчислить как:

$$\frac{\text{число ошибок, выявленных в проверенной выборке}}{\text{доля отобранных случаев в партии для проверки}}$$

К этому можно добавить число случаев ошибок в партии, проверенных на 100%-ной основе, с тем чтобы получить общее число случаев с ошибками. Если результат разделить на общее число закодированных записей (или пунктов), то получаем коэффициент входных ошибок.

2) Коэффициент выходных ошибок — это доля случаев, содержащих ошибки после исправления ошибок, выявленных в ходе контроля качества. Предполагается, что в проведенной выборке или партии, проверенной на 100%-ной основе, ошибок не остается. Число ошибок в непроверенной части партии получают умножением общего количества случаев в этой части на отношение:

число ошибок, установленных в проверенной части партии
общее число случаев в проверенной выборке из партии

Если число таких ошибок суммировать по всем (проверенным выборочно) партиям, то получим оценку общего числа остающихся ошибочных кодов; делением на общее число обрабатываемых случаев получаем коэффициент выходного качества. Сопоставление входного и выходного коэффициентов дает представление о непосредственном эффекте мер контроля качества с точки зрения сокращения доли ошибок. Проследив за коэффициентами выходных ошибок, можно получить представление о степени контроля за ошибками. Можно установить “предельный коэффициент выходных ошибок”; его превышение будет указывать на необходимость улучшить подготовку кодировщиков, контроль за их работой, методы кодирования и процедуры контроля качества.

3) Представляют практический интерес также показатели общего коэффициента проверки и коэффициент забракованных партий. Первый показатель — это общее число проверенных случаев (включая выборочную и 100%-ную проверку) как доля общего числа обработанных случаев. Очевидно его влияние на затраты по проверке и выходное качество. Второй показатель — это попросту доля бракованных партий, требующих 100%-ной проверки. Высокие доли указывают на плохое качество первоначального кодирования и/или слишком строгие критерии контроля качества.

В заключение следует еще раз подчеркнуть, что при обследованиях рабочей силы следует закодировать как можно больше вопросов, с тем чтобы уменьшить затраты на кодирование и связанные с ним ошибки.

Ввод данных

Речь идет об операции преобразования зарегистрированных или закодированных данных в машиночитаемые форматы. В ряде более развитых стран формы переписей и обследований считываются непосредственно на основе оптических знаков или распознавания символов. При таком методе основным источником ошибок ввода данных является невозможность считывать информацию из-за качества входных документов или оборудования. Описание процедуры и контроля качества при ней см., например, в работе (United States, 1978, p. 37—52).

Ввод данных оператором — это по-прежнему самый распространенный из применяемых методов. Процедуры измерения качества ввода данных операторами весьма схожи с теми, что описаны выше применительно к ручному кодированию.

Литература

- Canada. 1976. *Methodology of the Canadian Labour Force Survey*. Ottawa, Statistics Canada, Catalogue No. 71—526, Occasional.
- 1978. *A compendium of methods of error evaluation in censuses and surveys*. Ottawa, Statistics Canada. Catalogue No. 13—564, Occasional.
- 1982. *Labour force survey: Quality report* (See also accompanying Description of quality measures.) Ottawa, Statistics Canada.
- Dalenius, T. 1977. "Bibliography on non-sampling errors in surveys", in *International Statistical Review*, 45, pp. 71—89, 187—197, 303—317.
- Fellegi, I.; Sunter, A. 1974. Balance between different sources of survey errors: Some Canadian experiences", in *Sankhya, Indian Journal of Statistics*, 36 C (3), pp. 119—142.
- Gonzalez, M. et al. 1975. "Standards for discussion and presentation of errors in survey and census data", in *Journal of the American Statistical Association*, 70, No. 351, Pt. II.
- Hansen, M. H. et al. 1961. "Measurement of errors in censuses and surveys", in *Bulletin of the International Statistical Institute*, 38 (2), pp. 359—374.
- Hong Kong. 1980. *Report on March 1980 Labour Force Survey*. Hong Kong, Census and Statistics Department.
- International Labour Office (ILO), 1983. Thirteenth International Conference of Labour Statisticians: *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*. Geneva.
- 1986. *Statistical sources and methods*, Vol. 3. Economically active population, employment, unemployment and hours of work (household surveys). Geneva.
- Jabine, T. B.; Tepping, B. J. 1973. "Controlling the quality of occupation and industry data", in *Bulletin of the International Statistical Institute*, 45 (3), pp. 360—389.
- Kalton, G. 1977. "Practical methods for estimating survey sampling errors", in *Bulletin of the International Statistical Institute*, 47 (3), pp. 495—514.
- ; Blunden, R. 1973. "Sampling errors in the British General Household Survey", in *Bulletin of the International Statistical Institute*, 45 (3), pp. 83—97.

- Kish, L. 1962. "Studies in interviewer variance for attitudinal variables", in *Journal of the American Statistical Association*, 57, pp. 92—115.
- 1965. *Survey sampling*. New York, Wiley.
- ; Frankel, M. 1974. "Inference from complex samples", in *Journal of the Royal Statistical Society*, 13, 36, pp. 1—37.
- et al. 1976. *Sampling errors for fertility surveys*, WFS Occasional Papers No. 17, The Hague, International Statistical Institute.
- Lahiri, D. B. 1957. "Observations on the use of interpenetrating samples in India", in *Bulletin of the International Statistical Institute*, 36 (3), pp. 144—152.
- Lyberg, L. 1982. "Coding of occupation and industry: Some experiences from Statistics Sweden", in *Bulletin of Labour Statistics* (Geneva, ILO), 1982—3.
- Martin, J.; Butcher, B. 1982. "The quality of proxy information: Some results from a large-scale study", in *The Statistician*, 31 (4), pp. 293—319.
- Mellow, W.; Siden, H. 1983. "Accuracy of response in labour market surveys: Evidence and implications", in *Journal of Labour Economics*, Vol. 1 (4), pp. 331—344.
- Niemi, I. 1983. "Systematic bias in hours worked?", in *Statistisk Tidskrift*, (Stockholm, Statistics Sweden), 4, pp. 326—330.
- Schnell, D. et al. 1988. *PC CARP: Personal computer variance software for complex surveys*. Ames; Iowa State University (mimeographed).
- Sweden. 1980a. *A non-response study in the Swedish labour force surveys*, Methodological Studies No. 10E. Stockholm, Research Institute for Statistics on Living Conditions, National Central Bureau of Statistics.
- 1980b. *Response quality in the Swedish labour force surveys*, Methodological Studies No. 13E. Stockholm, Research Institute for Statistics on Living Conditions, National Central Bureau of Statistics.
- United Nations. 1982. *Non-sampling errors in household surveys*, NHSCP Technical Studies. New York, Department of Technical Co-operation for Development and UN Statistical Office.
- United States. 1978. *An error profile: employment as measured by the Current Population Survey*. Statistical Policy Working Paper No. 3. Washington, DC, Office of Federal Statistical Policy and Standards.
- Verma, V. 1980. "Sampling for national fertility surveys", in *Proceedings*, WFS Conference. The Hague, International Statistical Institute, 3, pp. 389—436.

- 1981. "Assessment of errors in household surveys", in *Bulletin of the International Statistical Institute*, 49 (2), pp. 905—919.
 - 1982. "The estimation and presentation of sampling errors", in *WFS Technical Bulletin*, No. 3, The Hague, International Statistical Institute.
 - ; Pearce, M. 1987. *Users' manual for CLUSTERS*, version 3.0. The Hague, International Statistical Institute.
 - et al. 1980. "Sample designs and sampling errors for the world fertility survey", in *Journal of the Royal Statistical Society A*, 143 (4), PP. 431—473.
- Wolter, K. 1985. *Introduction to variance estimation*, New York, Springer-Verlag.
- Woodruff, R.; Causey, B. 1976. "Computerised method for approximating the variance of a complicated estimate", in *Journal of the American Statistical Association*, 71, pp. 315—321.
- Zarkovich, S. S. 1979. "Stability of variance patterns", in *Journal of Indian Society of Agricultural Statistics*, 31 (1), pp. 23—48.

ПРИЛОЖЕНИЯ

Приложение I

РЕЗОЛЮЦИЯ, ОТНОСЯЩАЯСЯ К СТАТИСТИКЕ ЭКОНОМИЧЕСКИ АКТИВНОГО НАСЕЛЕНИЯ, ЗАНЯТОСТИ, БЕЗРАБОТИЦЕ И НЕПОЛНОЙ ЗАНЯТОСТИ, ПРИНЯТАЯ 13-Й МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИЕЙ СТАТИСТИКОВ ТРУДА (ОКТАБРЬ 1982 г.)

13-я Международная конференция статистиков труда,.....

.....
имея в виду существующие международные стандарты, относящиеся к занятости и безработице рабочей силы и содержащиеся в Резолюции I, принятой 8-й Конференцией (1954 г.), и относящиеся к измерению и анализу неполной занятости и неполному использованию трудовых ресурсов, содержащиеся в Резолюции III, принятой 11-й Конференцией (1966 г.);

признавая необходимость пересмотреть и расширить существующие стандарты с целью усилить их действенность в обеспечении технических ориентиров для всех стран, и особенно тех, где статистика менее развита, а также признавая важность таких стандартов в деле увеличения международной сопоставимости данных;

29 октября 1982 г. принимает следующую резолюцию взамен Резолюции I 8-й Конференции и параграфов 4—9 и 13 Резолюции III 11-й Конференции:

ЦЕЛИ И СФЕРА ДЕЙСТВИЯ

1. Каждая страна должна стремиться к разработке всеобъемлющей системы статистики экономической активности населения с целью обеспечить должную статистическую базу для различных пользователей данных с учетом специфических потребностей и обстоятельств данной страны. В частности, система должна удовлетворять потребности, связанные с измерением имеющихся и неиспользованных величин рабочего времени и трудовых ресурсов для целей макроэкономического анализа и планирования развития людских ресурсов, а также для измерения взаимосвязи между доходом от занятости и другими социально-экономическими характеристиками с целью формулирования и контроля за ходом выполнения политики и программ в области занятости, систем поддержания доходов, профессиональной подготовки и других аналогичных программ.

2. Для достижения указанных выше целей программа статистики экономически активного населения должна, в принципе, охватывать все сферы экономической деятельности, все сектора экономики и все группы занятых (наемных работников, лиц, занятых на принадлежащих им предприятиях, и т.д.) и должна быть разработана, по возможности, в увязке с другими видами экономической и социальной статистики. Программа призвана конкретно удовлетворять текущие и долгосрочные запросы, т.е. предусматривать данные для текущих целей, собираемые часто и на повторяющейся основе, и данные, получаемые через более длительные интервалы для глубокого структурного анализа и в качестве исходной информации:

а) программа текущей статистики должна отражать данные об экономически активном населении и его составных частях в данный момент таким образом, чтобы можно было надлежащим образом проследить тенденции и сезонные колебания. Как минимум страны должны собирать данные об экономически активном населении два раза в год, по возможности с учетом пика сельскохозяйственных работ и межсезонья, когда это считается желательным;

б) программа долгосрочной статистики, которая может включать переписи и обследования, должна обеспечивать: 1) исчерпывающие данные об экономически активном населении, 2) подробные данные о характере деятельности экономически активного населения в течение года, о взаимосвязи между занятостью, доходом и другими социально-экономическими характеристиками, 3) данные по другим вопросам (например, о детях и молодежи, домашних хозяйствах, возглавляемых женщинами), что определяется долгосрочными и постоянными потребностями.

3. Переписи населения и выборочные обследования домашних хозяйств и индивидов обычно составляют необходимые источники получения данных об экономически активном населении, которые можно увязать с данными по смежным вопросам. Обследования предприятий и учетные данные административных органов также могут служить источниками получения в ряде случаев более точных, более частых и более подробных данных о некоторых составляющих экономически активного населения. Различные источники информации необходимо рассматривать как дополняющие друг друга и их надо использовать в сочетании для получения, когда это необходимо, интегрированных совокупностей данных. Планируя переписи населения, обследования домашних хозяйств и индивидов или другие способы сбора данных об экономически активном населении, следует руководствоваться, насколько это возможно, международными стандартами.

4. С целью обеспечить сопоставимость данных между странами, где национальные понятия и определения неполностью соответствуют международным стандартам, необходимо давать пояснения, а основные агрегированные величины по возможности исчислять на основе как национальных, так и международных стандартов. Или же необходимые элементы агрегированных показателей следует приводить отдельно, с тем чтобы их можно было преобразовать из национальных стандартов в международные.

ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Экономически активное население

5. Экономически активное население включает всех лиц обоего пола, которые предлагают рабочую силу для производства экономических благ и услуг, как они определены в Системах национальных счетов и балансов ООН, в течение определенного отчетного периода. Согласно названным системам, производство экономических благ и услуг включает все производство и переработку первичных продуктов, будь то для рынка, для бартера или для собственного потребления; производство всех других благ и услуг для рынка и, в случае, когда домашние хозяйства производят такие блага и услуги для рынка, соответствующее производство для собственного потребления.

6. Двумя важными показателями экономически активного населения служат обычно активное население, учтенное за продолжительный отчетный период, например год, и население, активное в данный момент, или, что то же самое, "рабочая сила", учитываемая применительно к краткому отчетному периоду, например, неделе или дню.

Обычно активное население

7. (1) Обычно активное население включает всех лиц старше определенного возраста, которые большее число недель или дней в течение продолжительного периода (например, предыдущих 12 месяцев или предшествующего календарного года) были занятыми или безработными, как это определено в параграфах 9 и 10.

(2) Когда это понятие считают полезным и приемлемым, обычно активное население можно подразделить на занятых и безработных в соответствии с основным видом активности.

8. Рабочая сила, или “население, активное в данный момент”, включает всех лиц, удовлетворяющих требованиям для отнесения к числу занятых или безработных, которые содержатся в параграфах 9 и 10.

Занятость

9. (1) “Занятые” включают всех лиц старше определенного возраста, которые в течение определенного краткого периода, — либо одной недели, либо одного дня — относились к следующим категориям:

а) “наемные работники”:

а1) “работающие”: лица, которые в течение отчетного периода выполняли любую работу за заработную плату или жалованье деньгами или натурой;

а2) “имеющие рабочее место, но не работающие”: лица, которые уже ранее работали на своем рабочем месте, но временно не работали в течение отчетного периода, сохраняя формальную связь со своим рабочим местом.

Эту формальную связь с рабочим местом следует определять с учетом национальных особенностей и одного или нескольких из числа следующих критериев:

1) получение заработной платы или жалованья;

2) заверение в возможности вернуться на работу после того, как изменятся определенные обстоятельства, или наличие согласованной даты возвращения на работу;

3) истечение срока отсутствия на работе, которым, в определенных случаях, может быть тот срок, в течение которого работник может получать компенсационные выплаты без обязательства соглашаться на другую работу;

б) “занятые на собственном предприятии” (“самозанятые”):

б1) “работающие”: лица, которые на протяжении отчетного периода выполняли любую работу для получения прибыли или блага семьи, деньгами или натурой;

б2) “имеющие предприятие, но не работающие”: лица, имеющие предприятие — производственное, сельскохозяйственное или в сфере обслуживания, которые в течение отчетного периода временно не работали по любой конкретной причине.

(2) Для учетных целей понятие “любая работа” может означать работу по меньшей мере в течение одного часа.

(3) Лица, временно не работающие из-за болезни или травмы, праздника или отпуска, забастовки или локаута, учебного отпуска, рождения ребенка, свертывания экономической активности, временной дезорганизации или приостановки работы в связи с такими причинами, как плохая погода, механическая поломка, авария электросети, нехватка сырья или топлива либо по другой уважительной или неуважительной причине, должны относиться к наемным работникам при условии, что они сохраняют формальную связь с предприятием.

(4) Работодатели, лица, занятые на собственных предприятиях, и члены производственных кооперативов должны считаться занятыми на собственных предприятиях ("самозанятыми") и относиться к "работающим" и "неработающим" в зависимости от обстоятельств.

(5) Неоплачиваемых работающих членов семьи следует считать занятыми на собственных предприятиях ("самозанятыми") независимо от числа часов, отработанных за отчетный период. Страны, которые по тем или иным причинам предпочитают устанавливать минимальную продолжительность рабочего времени в качестве критерия для отнесения неоплачиваемых членов семьи к числу занятых, должны выявить и отдельно показать тех, кто работал меньше установленного времени.

(6) Лица, занятые производством экономических благ и услуг для собственного потребления или потребления в домашнем хозяйстве, должны относиться к работающим на собственных предприятиях, если такое производство обеспечивает существенный вклад в общее потребление домашнего хозяйства.

(7) Учеников, получающих заработную плату деньгами или натурой, необходимо считать наемными работниками и относить к "работающим" и "неработающим" точно так же, как других наемных работников.

(8) Студенты, домашние хозяйки и другие, занятые в течение отчетного периода в основном неэкономической деятельностью, но одновременно работавшие по найму или на собственном предприятии, как это определено в подпараграфе (1), должны относиться к занятым на тех же основаниях, что и другие категории занятых лиц и быть показаны отдельно, где это возможно.

(9) Военнослужащих следует относить к лицам наемного труда. К военнослужащим относят как регулярный, так и временный персонал, как это определено в последней резолюции Международной стандартной классификации профессий.

10. (1) “Безработные” включают всех лиц старше определенного возраста, которые в течение отчетного периода были:

а) “без работы”, т.е. не работали по найму или на собственном предприятии, как это определено в параграфе 9;

б) “готовы были в данный момент к работе, т.е. готовы были к наемному труду или работе на собственном предприятии в течение отчетного периода;

в) “искали работу”, т.е. предпринимали конкретные шаги в определенный прошлый период к поиску наемной работы или работы на собственном предприятии. Конкретные шаги могут включать регистрацию на государственной или частной бирже труда; обращение к работодателям; дежурство на стройках, фермах, у заводских ворот, на рынках или в других местах сбора; обращение за содействием к друзьям или родственникам; поиск земли, помещений, оборудования и машин для создания собственного предприятия; обращение за финансовыми ресурсами; обращение за разрешениями: лицензиями и т.п.

(2) В ситуациях, когда традиционные методы поиска работы лишены смысла, рынок труда в целом неорганизован или имеет ограниченные масштабы, когда спрос на рабочую силу недостаточен или она в основном занята на собственных предприятиях, стандартное определение, приведенное в подпараграфе (1), можно применять, ослабив критерий, относящийся к поиску работы.

(3) При применении критерия готовности приступить к работе в данный момент в ситуациях, предусмотренных подпараграфом (2), следует провести соответствующие тесты с учетом национальных особенностей. Эти тесты могут базироваться на таких понятиях, как желание работать в настоящее время или прошлый опыт работы, готовность приступить к работе за заработную плату или жалованье на преобладающих на месте условиях, или стремление открыть собственное дело при наличии необходимых ресурсов и условий.

(4) Несмотря на наличие критерия поиска работы, включенного в стандартное определение безработицы, лица, не имеющие работы и готовые приступить к ней, которые договорились начать работу по найму или на собственном предприятии после отчетного периода, должны считаться безработными.

(5) Лица, временно отсутствующие на своем рабочем месте и не имеющие формальной связи с предприятием, которые готовы приступить к работе и ищут ее, должны быть отнесены к безработным в соответствии со стандартным определением безработицы. Однако, с

учетом национальных условий и проводимой политики, страны могут ослабить критерий поиска работы применительно к временно уволенным лицам. В таких случаях временно уволенные лица, которые не ищут работу, но считаются безработными, должны быть выявлены и показаны отдельно.

(6) Студенты, домашние хозяйки и другие лица, которые в отчетном периоде были заняты в основном неэкономической деятельностью, но удовлетворяют критериям, предусмотренным в подпараграфах (1) и (2), должны считаться безработными на тех же основаниях, что и другие категории безработных лиц и, где это возможно, показаны отдельно.

Экономически неактивное население

11. “Экономически неактивное население” включает всех лиц, независимо от возраста, и в том числе тех, кто моложе возраста, определенного для учета экономически активного населения, как это предусмотрено параграфом 5.

Население, не активное в данный момент

12. (1) “Население, не активное в данный момент”, или, что то же самое, лица вне рабочей силы, включает всех индивидов, которые не были занятыми или безработными в течение данного короткого периода, а следовательно, не были активны в данный момент из-за а) посещения учебных заведений, б) выполнения домашних обязанностей, в) пожилого возраста или получения пенсии по старости, г) других причин, например, болезни или инвалидности, что следует уточнить.

(2) Страны, которые придерживаются стандартного определения безработицы, могут выявить лиц, не отнесенных к безработным, которые были готовы приступить к работе, но не искали ее в отчетном периоде и показать их отдельно в составе населения, не активного в данный момент.

Обычно неактивное население

13. (1) “Обычно неактивное население” включает всех лиц, чей основной статус в течение определенного продолжительного периода не позволяет отнести их ни к занятым, ни к безработным. К нему относятся следующие категории лиц: а) учащиеся и студенты, б) домашние хозяйки, в) получатели доходов (пенсионеры, рантье и

т.п.); г) прочие (получатели государственной или частной помощи, дети, не посещающие школу и т.п.), как это определено в “Принципах и рекомендациях по проведению переписей населения и жилья” ООН (1980 г.).

(2) Где это необходимо, можно включить дополнительные категории, имея в виду выявить: 1) лиц, занятых на неоплачиваемых общественных или добровольных работах; 2) других лиц, занятых различными видами деятельности, не подпадающих под определение экономических.

Неполная занятость

14. Неполная занятость имеет место, когда работа индивида неудовлетворительна с точки зрения определенных нормативов или другой возможной работы с учетом его квалификации (подготовки и опыта работы). Можно выделить две основные формы неполной занятости: видимую и невидимую.

15. (1) Видимая неполная занятость — это преимущественно статистическое понятие, которое поддается непосредственному измерению в рамках обследований рабочей силы и других обследований и отражает недостаточный объем занятости.

(2) Невидимая неполная занятость — это преимущественно аналитическое понятие, отражающее неправильное распределение трудовых ресурсов или глубокий дисбаланс между трудом и другими факторами производства. Характерными симптомами могут служить низкие доходы, неполное использование квалификационных возможностей, низкая производительность. Аналитические исследования невидимой неполной занятости следует направлять на изучение и анализ широкого круга данных, включая доход и уровень квалификации (скрытая неполная занятость) и показатели производительности (потенциальная неполная занятость).

16. По практическим соображениям статистическое изучение неполной занятости можно ограничить ее видимым аспектом.

Видимая неполная занятость

17. Следует различать два элемента измерения видимой неполной занятости:

- а) число лиц, занятых неполное рабочее время;
- б) объем (масштабы) видимой неполной занятости.

18. (1) Лица, занятые неполное рабочее время, включают всех тех, кто работает по найму или на собственном предприятии, работающих или не работающих в данный момент, которые в отчетном периоде не по своей воле были заняты меньше нормального времени, установленного для данного вида деятельности, искали дополнительную работу и были готовы приступить к ней.

(2) Для целей отнесения лиц к занятым неполное время нормальную продолжительность рабочего времени для данного вида деятельности следует принимать с учетом национальных особенностей, отраженных в законодательстве страны, если это применимо, и обычной практики — в других случаях, либо на основе единой общепринятой нормы.

Объем видимой неполной занятости

19. (1) Объем видимой неполной занятости можно измерить путем агрегирования времени возможной дополнительной работы каждого лица, занятого неполное время в отчетном периоде. Возможное дополнительное рабочее время можно измерять в рабочих днях, половинах дней или часах в зависимости от национальных условий и характера собранных данных. Может оказаться целесообразным отдельно измерить ту часть объема видимой неполной занятости, которая соответствует “потерянному времени”, определяемому как разность между обычно отработываемым и фактически отработанным временем.

(2) Страны, которые хотят применить критерий поиска работы для целей измерения объема видимой неполной занятости, могут сделать это, приняв во внимание продолжительность искомой работы.

20. Итоговую оценку объема текущей безработицы и видимой неполной занятости можно получить на основе распределения рабочего времени всех лиц в рабочей силе с учетом общего времени каждого лица, потенциально возможного для трудовой деятельности, и фактически отработанного времени. Этот показатель можно, ради простоты, измерить в рабочих днях, половинах дней или же, более полно, в часах, когда это возможно.

Аналитические понятия

21. С учетом понятий и определений, содержащихся в параграфах 5 — 20, можно получить разнообразные аналитические понятия и показатели. Например:

(1) Экономически активное население можно разделить на две основные группы: вооруженные силы и экономически активное гражданское население.

(2) Экономически активное население можно соотнести с общим населением, чтобы получить показатель участия в рабочей силе, или, что точнее, соотнести с населением старше возраста, установленного для учета экономически активного населения.

(3) Занятое население можно соотнести с общим населением старше определенного возраста, чтобы получить коэффициент занятости населения.

(4) Безработное население можно соотнести с экономически активным населением, чтобы получить общий коэффициент безработицы. Где это целесообразно и возможно, коэффициенты безработицы можно исчислить отдельно для наемных работников и занятых на собственных предприятиях.

(5) Число лиц, занятых неполное рабочее время, можно соотнести с занятыми лицами, а соответствующие доли можно исчислить отдельно по отраслям экономической деятельности и профессиональным группам.

(6) Совокупный уровень безработицы и неполной занятости исчисляют как отношение неиспользованного рабочего времени, потенциально возможного для работы, к фактическому рабочему времени или возможному рабочему времени. Перечисленные выше доли и отношения можно рассчитать отдельно для лиц обоего пола и определенных возрастных групп.

22. Метод, основанный на распределении рабочего времени и предложенный в параграфе 20, реализуют на основе ряда текущих обследований, которые охватывают представительную выборку отчетных периодов на протяжении года, и в этом случае его можно применить для оценки использованного и неиспользованного времени в течение года. Оценки можно выразить в человеко-днях или человеко-часах или же, если желательно, преобразовать в стандартные полные человеко-годы.

Взаимосвязь между занятостью и доходами

23. Для достижения целей анализа взаимосвязи между занятостью и доходом, упомянутых в параграфе 1, странам необходимо разработать программы сбора данных о занятости и доходах, отражающих соответствующие экономические и социальные аспекты. В частности, следует получить данные о занятости, доходе от занятости и доходах домашних хозяйств с целью: а) проанализировать

способность различных видов экономической деятельности приносить доход, б) выявить число и характеристики лиц, не способных обеспечить свое экономическое благосостояние на основе доступных им возможностей для трудоустройства.

24. (1) С целью отразить всесторонние взаимосвязи между занятостью и доходом, данные о занятости, доходе от занятости и доходе домашних хозяйств должны относиться к работе населения в течение продолжительного периода, желательно года, с учетом не только основного занятия, но и дополнительных видов деятельности и источников дохода.

(2) Доход от занятости включает заработную плату, жалованье и другие заработки деньгами и натурой лиц наемного труда и чистый предпринимательский доход лиц, работающих на собственных предприятиях.

(3) Понятия и определения дохода и его элементов содержатся в резолюциях, относящихся к интегрированной системе статистики заработной платы и к обследованиям доходов и расходов домашних хозяйств, принятых 12-й Конференцией (1973 г.), а также во "Временных рекомендациях по статистике распределения дохода, потребления и накопления домашних хозяйств" ООН (1977 г.).

(4) Данные о занятости и доходе следует анализировать, насколько это возможно, в увязке с продолжительностью рабочего времени, размером домашнего хозяйства, числом работающих, активами и другими демографическими, социальными и экономическими характеристиками индивида и домашнего хозяйства.

(5) Данные о занятости и доходе должны быть совместимы и, насколько это возможно, интегрированы с показателями экономически активного населения, перечисленными выше в параграфах 5— 22.

СБОР, АНАЛИЗ И ГРУППИРОВКА ДАННЫХ

25. Международное бюро труда должно подготовить руководство по статистике экономически активного населения, занятости, безработицы и неполной занятости с детальным описанием методологии сбора, табулирования и анализа данных.

26. (1) Анализ экономически активного населения и экономически неактивного населения должен включать группировку в зависимости от важных демографических, социальных и экономических характеристик, равно как и комбинационные группировки по двум или нескольким взаимосвязанным признакам.

(2) В частности, население старше возраста, установленного для учета экономически активного населения, следует сгруппировать в зависимости от обычного статуса активности (занятые и безработные лица, учащиеся, домохозяйки и т.д.) и статуса в данный момент (занятые, безработные и экономически не активные в данный момент).

27. Для обеспечения международных сопоставлений группировки данных об экономически активном населении должны отражать классификации или поддаваться преобразованию в недавно принятые международные стандартные классификации, такие как:

а) Международная стандартная классификация профессий (ISCO) — МОТ;

б) Международная стандартная классификация всех видов экономической деятельности (ISIC) — ООН;

в) международные классификации в зависимости от статуса (наемный работник, работодатель и т.д. — определения статуса даны ООН — за исключением того, что при отнесении к неоплачиваемым работающим членам семьи минимальный критерий времени (по меньшей мере треть нормального рабочего времени) более применять не следует;

г) Временные рекомендации по стандартной международной классификации возрастов — ООН.

28. Для группировок по другим признакам, например, продолжительность рабочего времени или длительности безработицы, Международное бюро труда должно разработать соответствующие международные стандартные классификации с учетом существующей практики и потребности отдельных стран.

ДАННЫЕ ПО ОТДЕЛЬНЫМ ВОПРОСАМ

29. С целью должного изучения переходных этапов от учебы к трудовой деятельности и разработки в необходимых случаях соответствующих мер следует периодически получать определенные данные о детях и молодежи с точки зрения посещения ими учебных заведений и участия в экономической деятельности. Для указанной цели может быть необходимо собирать дополнительные сведения о детях и молодых людях моложе минимального возраста, принятого для учета экономически активного населения.

30. (1) Для разработки и контроля за реализацией программ, относящихся к участию женщин в экономическом развитии и стимулированию равенства полов, существенно иметь надлежащую статистическую базу. Поэтому данные об экономически активном

населении, занятости, безработице, неполной занятости и по смежным вопросам необходимо собирать отдельно по мужчинам и женщинам.

(2) Кроме того, для получения более точных данных об участии женщин в экономической деятельности методы учета следует тщательно проверить, чтобы избежать погрешностей в охвате мужчин и женщин. Погрешности в виде недооценки участия женщин в экономической деятельности могут быть результатом, например, неполного охвата неоплачиваемых видов экономической деятельности, неспособности респондентов и счетчиков выявить многообразие деятельности женщин и использования ответов других лиц. Когда это необходимо, следует провести исследование с целью выявить масштабы, характер и источники возможных погрешностей, если они имеются, и выработать меры по их недопущению.

31. Поскольку участие индивидов в экономической деятельности часто зависит от других членов семьи или домашнего хозяйства, а во многих странах, особенно в сельской местности развивающихся стран, экономическая жизнь во многом организована на семейной основе или в рамках домашнего хозяйства, данные об экономически активном населении, занятости, безработице, неполной занятости и по другим вопросам следует периодически дополнять данными о семьях и домашних хозяйствах: например, указывать, как безработные соотносятся с другими членами домашнего хозяйства или семьи; наличие других работающих членов домашнего хозяйства или семьи, количество детей в домашнем хозяйстве или семье, а также общее число безработных в домашнем хозяйстве или семье; пол и другие характеристики основного кормильца в семье и т.п.

32. Для получения более точной и подробной информации о занятости, безработице и неполной занятости, а также для других целей, например, для выявления многообразия деятельности и второстепенных ее видов, необходимо попытаться периодически получать данные об использовании времени.

33. С целью учесть деятельность неформального сектора в развивающихся и развивающихся странах, а также несельскохозяйственные виды деятельности, обычно осуществляемые домашними хозяйствами в сочетании с сельскохозяйственной деятельностью в развивающихся странах, а также учитывая ограниченность информации по этим вопросам, желательно, чтобы страны разработали соответствующие методологии и программы сбора данных о городском неформальном секторе и сельских несельскохозяйственных видах деятельности. В частности, следует выработать определения и классификации, которые позволили бы выявить и сгруппировать эконо-

мически активное население в городском неформальном секторе, а также лиц, занятых несельскохозяйственной деятельностью в сельской местности.

34. С целью обеспечить соответствующие возможности для трудоустройства и получение дохода для инвалидов и других нетрудоспособных лиц необходимо, используя обоснованную методологию, собрать данные о количестве таких лиц и их распределении с учетом социальных и экономических характеристик, выделив, в частности, занятых, безработных и экономически неактивных.

35. (1) Рекомендуются, чтобы в странах с планируемой экономикой широко использовались балансы рабочей силы, с тем чтобы определить масштабы и структуру рабочей силы и ее географическое распределение в зависимости от характера занятости и сектора национальной экономики.

(2) Население трудоспособных возрастов, за исключением неработающих инвалидов, а также население нетрудоспособных возрастов, считаются трудовыми ресурсами. Баланс трудовых ресурсов можно разработать отдельно для мужчин и женщин, выявив лиц, занятых вспомогательным фермерством или домашней работой, неработающих инвалидов трудоспособного возраста и лиц нетрудоспособного возраста.

(3) Данные баланса трудовых ресурсов позволяют выявить долю лиц, которых можно будет использовать в будущем для работы в национальной экономике.

36. Предполагается, чтобы страны собирали информацию об экономически неактивном населении с учетом национальных потребностей и условий, с тем чтобы содействовать правительствам в планировании людских ресурсов и выработке политики развития. Необходимо разработать классификации, которые обеспечили бы составление комбинационных группировок, отражающих степень принадлежности к рынку труда лиц, указанных выше в параграфах 12 (1) и 13 (1).

ОЦЕНКА И РАСПРОСТРАНЕНИЕ ИНФОРМАЦИИ

37. Подобно любым другим данным, информация об экономически активном населении, занятости, безработице, неполной занятости и по смежным вопросам подвержена ошибкам. Даже когда программа сбора данных разработана весьма тщательно с целью минимизации возможных ошибок, некоторое их число будет присутствовать. Поэтому надлежащая интерпретация результатов требует знания качества данных. Оценка качества данных необходима

также для улучшения процедур сбора, обработки и оценки данных при последующих раундах программы. Процедура оценки должна, по возможности, быть составной частью программы сбора данных.

38. Данные об экономически активном населении, занятости, безработице, неполной занятости и смежных вопросах следует публиковать быстро и сделать их широко доступными. Данные можно обнародовать этапами путем издания предварительных отчетов по мере получения сводных показателей, а затем дополнить окончательными отчетами с уточненными и подробными данными в табличной форме и, где это возможно и допустимо, — в машиночитаемой форме.

39. Каждый выпуск данных об экономически активном населении, занятости, безработице, неполной занятости и смежных вопросах должен содержать четкие указания на характер данных и их подробное техническое описание. В частности, следует охарактеризовать охват данных, понятия и определения, метод сбора данных, объем и план выборки, если использован выборочный метод; процедуры оценки и поправки, включая сезонные поправки; показатели качества данных, в том числе ошибки выборки и регистрации, где это возможно, а также изменения во временных рядах; отклонения от международных стандартов и взаимосвязь с другими источниками аналогичных данных и совокупностями данных.

**РЕЗОЛЮЦИЯ, ОТНОСЯЩАЯСЯ К ИЗМЕРЕНИЮ
И АНАЛИЗУ НЕПОЛНОЙ ЗАНЯТОСТИ
И НЕПОЛНОГО ИСПОЛЬЗОВАНИЯ РАБОЧЕЙ СИЛЫ,
ПРИНЯТАЯ 11-Й МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИЕЙ
СТАТИСТИКОВ ТРУДА (ОКТАБРЬ 1966 г.)**

Извлечения

11-я Международная конференция статистиков труда,.....

.....
полагая, что уточненные ориентиры для дальнейшего развития методов измерения и анализа неполной занятости будут полезны, особенно в развивающихся странах, 26 октября 1966 г. принимает следующую резолюцию взамен Резолюции III 9-й Международной конференции статистиков труда:

Общие цели

1. Основная цель измерения и анализа неполной занятости и других аспектов неполного использования трудовых ресурсов заключается в содействии выработке и оценке краткосрочной и долгосрочной политики и мер, в частности, в области планирования и прогнозирования трудовых ресурсов для стимулирования "полной, производительной и свободно избираемой" занятости, как это определено Конвенцией и Рекомендацией (№122), относящимися к политике занятости и принятыми Международной конференцией по труду в 1964 г.

2. Особое внимание следует уделить неполной занятости в секторах экономики, регионах и применительно к категориям трудящихся, на которых особенно отражается неполная занятость и которые порождают острые проблемы в условиях данной страны. Примерами могут служить крестьяне-фермеры, особенно в развивающихся странах; мелкие предприятия, экономически отсталые районы, приходящие в упадок отрасли, например, угольная промышленность в индустриально развитых странах, сезонные работы, например в сельском хозяйстве или на строительстве, а также категории работников, особенно подверженных дискриминации в области занятости по признаку пола, возраста, национальности, расы и т.д.

3. В развивающихся странах предварительную информацию о неполной занятости можно получить в рамках обследований домашних хозяйств. Повторные или детальные обследования неполной занятости будут необходимы, когда желательно разработать или оценить краткосрочные или долгосрочные программы преодоления неполной занятости.

.....

Элементы и методы измерения и анализа невидимой неполной занятости

Скрытая неполная занятость

10. (1) Для анализа скрытой неполной занятости необходима информация о доходах. В развивающихся странах удовлетворительные оценки данных о доходах обычно можно получить в рамках выборочных обследований рабочей силы лишь применительно к наемным работникам; для других категорий работающих приемлемые данные о группах по размеру доходов можно получить с помощью тщательно проведенных обследований семейных бюджетов. В развитых странах, а в некоторых случаях в развивающихся странах удовлетворительные данные о заработках можно получить на основе выборочных обследований рабочей силы и других источников, например, налоговых отчетов.

(2) Методологию анализа скрытой неполной занятости в виде неполного использования квалификации работников все еще предстоит разработать. Для этой цели следует провести экспериментальные обследования и исследования, в частности, применительно к некоторым ограниченным категориям работников, которых затрагивает эта проблема.

Потенциальная неполная занятость

11. (1) Потенциальная неполная занятость как аспект неполной занятости, которую следует оценивать согласно критерию низкой производительности труда, существует, как можно полагать, в тех случаях, когда индивид работает на предприятии или в экономической единице, где производительность ненормально низка.

(2) Основное внимание при подходе к неполной занятости на основе критерия производительности концентрируется на динамической оценке взаимосвязи между предложением рабочей силы и динамикой потребности в ней по мере роста производительности в процессе экономического развития, а также на сопоставлениях производительности в отраслях и регионах.

(3) Потенциально неполностью занятые не могут быть выявлены непосредственно, но там, где имеются подробные данные о рабочей силе и производстве, анализ, основанный на этих данных, способен выявить общие показатели числа и характеристик лиц, потенциально неполностью занятых, а также объем такой неполной занятости.

(4) В различных секторах экономики, в частности, в сельском хозяйстве, оценки “излишков рабочей силы” или “резервов рабочей силы” можно получить путем сравнения имеющихся трудовых единиц и фактически используемых или требуемых единиц при различных допущениях относительно производительности.

Дальнейшие действия

12. Странам, проводящим исследования и анализ различных форм невидимой неполной занятости, особенно в отношении сельского хозяйства в развивающихся странах, рекомендуется сообщать о результатах Международному бюро труда, с тем чтобы соответствующие методы исследования были рассмотрены на будущей Международной конференции статистиков труда.

Статистическая и аналитическая работа

14. Следует рассмотреть вопрос об использовании и совершенствовании системы выборочных обследований рабочей силы в развивающихся странах, где это возможно. Такая система может быть необходима в качестве составной части всеобъемлющих данных о рабочей силе, требующихся, помимо прочего, для анализа неполной занятости. В то же время она может дать непосредственные оценки основных элементов неполной занятости, и особенно видимой неполной занятости.

15. Широкие обследования общего характера, такие как выборочные обследования рабочей силы, необходимо дополнить соответствующей программой ограниченных по масштабу, но специальных обследований, направленных на глубокое изучение неполной занятости или получение региональных и местных данных. Кроме того, рекомендуется, чтобы страны проводили методологические исследования с целью снижения неопределенности, присущей выборочным обследованиям, и изучили подходящие аналитические методы, обеспечивающие получение как можно более достоверных результатов.

16. Для измерения и анализа различных аспектов неполной занятости необходимо в должной мере воспользоваться существующими статистическими и техническими данными. К основным пригодным для этой цели статистическим источникам относятся, помимо выборочных обследований рабочей силы, обследования семейных бюджетов, переписи населения, сельскохозяйственные и промышленные переписи, периодические отчеты предприятий, отчеты служб трудоустройства и национальные счета. Источники данных включают специальные обследования и отчеты преимущественно нестатистического характера, например, обследования управления и использования рабочего времени на фермах и отчеты о программах повышения квалификации на селе.

17. Где это возможно, страны должны стремиться разрабатывать соответствующие программы статистических и аналитических исследований неполной занятости, ориентированных на текущие и настоятельные проблемы, особенно на потребности органов планирования развития.

Неполное использование трудовых ресурсов

18. Помимо лиц, которые в данный момент входят в состав рабочей силы, но вклад которых в доход семей и национальный продукт ограничен в результате безработицы и неполной занятости, во многих странах имеются индивиды, которые не принадлежат к рабочей силе, но вошли бы в нее при определенных обстоятельствах. Такие лица могут и не искать активно работу, например, из-за отсутствия подходящей для них работы; или же они могут иметь физический или умственный недостатки, которые можно было бы преодолеть с помощью учебы или других корректирующих мер. Хотя любой стране важно знать масштабы неполного использования ее трудовых ресурсов, удовлетворительные методы его измерения пока не разработаны и не опробованы. Странам, которые проводят изучение неполного использования трудовых ресурсов, предлагается сообщать о своем опыте Международному бюро труда, для того чтобы эффективные методы изучения можно было рассмотреть на будущей Международной конференции статистиков труда.

**РЕЗОЛЮЦИЯ, ОТНОСЯЩАЯСЯ К СТАТИСТИКЕ
РАБОЧЕГО ВРЕМЕНИ, ПРИНЯТАЯ
10-Й МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИЕЙ
СТАТИСТИКОВ ТРУДА (ОКТАБРЬ 1962 г.)**

И з в л е ч е н и я

Определения

4. (1) Нормальное рабочее время — это рабочее время, установленное законом или правилами либо в соответствии с ними, а также коллективными договорами и решениями арбитража.

(2) Там, где продолжительность рабочего времени не установлена законом или в соответствии с ним, а также коллективными соглашениями или решениями арбитража, нормальным следует считать то количество часов в день или неделю, сверх которого работа оплачивается по ставке сверхурочных работ или считается исключением из правила или обычая на данном предприятии применительно к определенным категориям трудящихся.

5. (1) Данные о фактическом рабочем времени должны включать:

а) фактически отработанные часы в течение нормальных периодов времени;

б) время, отработанное сверх часов в нормальные периоды времени и обычно оплачиваемое по более высокому, нежели нормальные, ставкам (сверхурочные);

в) время, затраченное на рабочем месте, например, для подготовки рабочего места, ремонта и обслуживания, подготовки и чистки инструментов, подготовки накладных, графиков и отчетов;

2) время, затраченное на рабочем месте по таким причинам, как отсутствие работы, поломка машин, несчастный случай, а также время, затраченное на рабочем месте без работы, но подлежащее оплате в соответствии с гарантированным трудовым договором;

д) время, соответствующее кратким периодам отдыха на рабочем месте, включая перерывы для чая или кофе.

(2) Из данных об отработанном времени исключают:

а) оплаченное, но не отработанное время, например, ежегодный отпуск, оплачиваемые государственные праздники, оплачиваемое время болезни;

б) перерывы на прием пищи;

в) время, затраченное на поездку на работу и с работы.

6. В связи со значительными различиями между странами в отношении оплаты праздничных дней и других периодов, когда работа не выполняется, представляется нецелесообразным на данном этапе принимать международные определения оплаченного времени. Однако многие страны придут к выводу, что данные об оплаченном времени хотя и не заменяют фактически отработанное время, все же могут быть полезны для внутренних целей и их обычно легко получить на основе ведомостей заработной платы и других бухгалтерских записей.

**РЕЗОЛЮЦИЯ, ОТНОСЯЩАЯСЯ К ИНТЕГРИРОВАННОЙ
СИСТЕМЕ СТАТИСТИКИ ЗАРАБОТНОЙ ПЛАТЫ,
ПРИНЯТАЯ 12-Й МЕЖДУНАРОДНОЙ КОНФЕРЕНЦИЕЙ
СТАТИСТИКОВ ТРУДА (ОКТАБРЬ 1973 г.)**

Извлечения

Зарботки

8. Понятие зарботков, в том виде, в каком его используют в статистике зарботной платы, относится к вознаграждению деньгами или натурой, выплачиваемому наемным работникам, как правило, через регулярные промежутки времени за отработанное время или проделанную работу, а также к выплатам за неотработанное время, например, за ежегодный отпуск, другие оплачиваемые отпуска и праздники. В зарботки не включаются взносы работодателей на социальное обеспечение своих работников, в пенсионные фонды, а также суммы, полученные работниками в рамках этих систем и фондов. В зарботках не учитывают также выходное пособие.

9. Данные о зарботках относятся к валовому вознаграждению работников, т.е. общей сумме до каких-либо вычетов, которые работодатель делает в связи с налогообложением, взносами работников в фонды социального обеспечения и пенсионные фонды, страхованием жизни, профсоюзными взносами и другими обязательствами работников.

10. (1) В зарботки включают: прямую зарботную плату и жалование, выплаты за нерабочее время (исключая выходное пособие), премии и подарки, а также субсидии на жилье и семейные пособия, выплаченные работодателем непосредственно работнику.

а) Прямая зарботная плата за отработанное время или проделанную работу включает: 1) выплаты зарботной платы повременщикам за отработанное время; 2) стимулирующие доплаты повременщиков; 3) зарботки сдельщиков (без доплаты за сверхурочные часы); 4) доплаты за работу в сверхурочные часы, ночное время, сменный график и за праздничные дни; 5) комиссионные, уплаченные торговому и иному персоналу. Также включаются: доплаты за стаж или особую квалификацию, районные надбавки, доплаты за

ответственность, грязные, опасные и дискомфортные условия труда, выплаты в соответствии с системами гарантированной заработной платы; надбавки в связи со стоимостью жизни и другие регулярно выплачиваемые надбавки.

б) Выплаты за нерабочее время включают оплачиваемые государственные праздники, ежегодные отпуска и другое нерабочее время с сохранением заработной платы.

в) Премии и подарки включают сезонные премии и премии за год, дополнительные выплаты к отпуску (сверх нормальных отпускных сумм) и премии в порядке распределения прибыли.

(2) В статистике заработков следует проводить различие между денежными заработками и выплатами натурой.

ОГЛАВЛЕНИЕ

Часть вторая. МЕТОДЫ	5
Глава 10. Планирование, организация и совершенствование обследований	5
1. Введение	5
2. Источники данных об экономически активном населении	6
3. Структура и организация обследований	12
4. Основные проблемы планирования обследований	30
5. Реорганизация обследований	36
Литература	41
Глава 11. Планирование выборки	43
1. Введение	43
2. Определение обследуемого населения	44
3. Основа выборки и использование главной выборки	50
4. Некоторые практические аспекты планирования выборки	66
5. Отбор во времени	75
6. Процедуры оценки	92
7. Изменение плана выборки	106
Литература	109
Глава 12. Разработка и построение анкеты	112
1. Введение	112
2. Уточнение содержания и результатов обследования	116
3. Некоторые примеры из национальных обследований рабочей силы	125
4. Принципы и методы построения анкеты	147
5. Реализация анкеты	161
Литература	165
Глава 13. Проведение обследования и обработка данных	167
1. Введение	167
2. Набор и учеба персонала	169
3. Опробование инструментов и процедур обследования	173
4. Основа выборки и формирование выборки	174
5. Сбор данных	175
6. Подготовка и обработка данных	180
7. Табулирование, отчетность и рассылка	192
Литература	198
Глава 14. Точность и оценка данных	199
1. Принципы выявления ошибок	199

2. Ошибки охвата и смежные ошибки	211
3. Неполучение ответов	219
4. Ошибка выборки	228
5. Ошибки в ответах	235
6. Ошибки при обработке	246
Литература	250
П р и л о ж е н и я	253
<i>Приложение 1. Резолюция, относящаяся к статистике экономически активного населения, занятости, безработице и неполной занятости, принятая 13-й Международной конференцией статистиков труда (октябрь 1982 г.)</i>	<i>253</i>
<i>Приложение 2. Резолюция, относящаяся к измерению и анализу неполной занятости и неполного использования рабочей силы, принятая 11-й Международной конференцией статистиков труда (октябрь 1966 г.)</i>	<i>268</i>
<i>Приложение 3. Резолюция, относящаяся к статистике рабочего времени, принятая 10-й Международной конференцией статистиков труда (октябрь 1962 г.)</i>	<i>272</i>
<i>Приложение 4. Резолюция, относящаяся к интегрированной системе статистики заработной платы, принятая 12-й Международной конференцией статистиков труда (октябрь 1973 г.)</i>	<i>274</i>

- Хуссманнс Р., Мехран Ф., Верма В.**
Х 24 **Обследования экономически активного населения: занятость, безработица и неполная занятость: Методологическое руководство МБТ: Пер. с англ. В 2 кн./Кн. 2 — М.: АО "Финстатинформ", 1994. — 278 с.**
ISBN 5-7166-0077-8

В книге подробно разъясняются международные стандарты учета занятости населения и безработицы, даны детальные классификации трудовых ресурсов, всесторонне рассматриваются различные категории занятых, безработных, занятых неполный рабочий день и критерии отнесения к той или иной группе.

Книга рассчитана на работников статистических органов, специалистов службы занятости, миграционной службы, экономистов, статистиков, демографов, для которых книга будет весьма полезна, учитывая целесообразность приближения отечественной информационно-статистической системы к международным стандартам учета и отчетности.

ББК 65.524.0я2

Хуссманнс Ральф, Мехран Фарад, Верма Виджай
ОБСЛЕДОВАНИЯ ЭКОНОМИЧЕСКИ АКТИВНОГО НАСЕЛЕНИЯ:
ЗАНЯТОСТЬ, БЕЗРАБОТИЦА И НЕПОЛНАЯ ЗАНЯТОСТЬ
(В 2 книгах)
КНИГА 2

Зав. редакцией О.А. Дмитриева
Редактор О.Н. Толыпина
Обложка художника Ф.Г. Миллера
Технический редактор М.М. Егорова
Корректор В.Б. Романова

Лицензия ЛР № 040559 от 17.12.92.

Подписано в печать 14.07.94. Формат 60х88/16. Бумага офсетная. Гарнитура "Таймс". Усл. печ. л. 17,15. Тираж 1500 экз. Заказ № 700

АО "Финстатинформ". Почтовый адрес: 103045, Москва, а/я 49
Телефоны: (095) 255-11-75, 255-65-69, факс: (095) 255-63-75
