

Distr.
GENERAL

CES/SEM.40/20
7 September 1998

RUSSIAN
Original: ENGLISH

**СТАТИСТИЧЕСКАЯ КОМИССИЯ и
ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ
КОМИССИЯ**

**СТАТИСТИЧЕСКОЕ УПРАВЛЕНИЕ
ЕВРОПЕЙСКИХ СООБЩЕСТВ (ЕВРОСТАТ)**

КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ

Совместная рабочая сессия ЕЭК-Евростат
по переписям населения и жилищ 1/
(Дублин, Ирландия, 9-11 ноября 1998 года)

Тема для исследования № 1

ПЕРЕПИСЬ НАСЕЛЕНИЯ 2000 ГОДА НА ОСНОВЕ АДМИНИСТРАТИВНЫХ ИСТОЧНИКОВ ДАННЫХ, КАЧЕСТВЕННЫЕ АСПЕКТЫ

Вспомогательный документ, представленный Статистическим
управлением Норвегии 2/

1. Введение

1. Перепись населения Норвегии в 2000 году будет проводиться полностью на основе данных регистров, в то время как перепись жилищ будет осуществлена в форме традиционного обследования. В настоящем документе рассматриваются вопросы, касающиеся переписи населения. Информация о переписи жилищ приводится в документе "Создание регистра жилья в Норвегии: недостающее звено".

1/ Документы, подготовленные для этой рабочей сессии, будут рассматриваться точно так же, как и документы, подготавливаемые для семинаров.

2/ Подготовлен Коэном Хендриксом и Паулем Инге Северайде, отдел переписей населения и жилищ.

GE.98-32271 (R)

2. Используемые Статистическим управлением Норвегии (СУН) регистры представляют собой так называемые статистические регистры. Входные данные для всех рассматриваемых статистических регистров берутся либо из административных регистров, либо из других административных источников.

3. Регистры используются СУН для сбора данных переписей начиная с 1970 года 3/. В ходе переписи 1990 года информация о демографических вопросах и об образовании, а также о географических характеристиках собиралась на основе регистров. Среди экономических показателей на основе регистров собирались данные о доходах. Данные о рынке труда были единственной темой переписи населения, в отношении которой проводилось обследование. Что касается лиц, проживающих в небольших районах, данные о рынке труда собирались исключительно на основе вопросников. В крупных районах проводилось выборочное обследование; полученные в результате этих обследований данные объединялись с данными регистров.

4. Во время переписи населения 2000 года СУН решила использовать административные данные для всех переменных, включая характеристики рынка труда. Такое решение объясняется улучшением в некотором отношении качества показателей после переписи 1990 года. Краткое описание норвежской системы данных регистров приводится в главе 2.

5. Однако основной темой настоящего документа являются качественные аспекты данных переписи, получаемых на основе регистров. Качественные аспекты характеризуются четырьмя признаками:

- улучшение качества;
- согласованность данных, получаемых из различных источников;
- разработка методов измерения качества;
- обследование вопросов контроля качества.

3/ Краткая история использования регистров при проведении переписей в Норвегии приводится в документе "Создание регистра жилья в Норвегии: недостающее звено".

2. Краткое описание норвежской системы данных регистров

Регистр населения

6. Наличие Центрального регистра населения (ЦРН), в котором используется специальный национальный идентификационный номер, имеет очень важное значение для сбора статистических данных на основе регистров в Норвегии. В настоящее время этот персональный идентификационный номер используется в регистрах всех уровней. Он служит также связующим звеном для согласования различных регистров. ЦРН является центральным регистром при проведении переписей населения.

Регистры видов деятельности и доходов

7. В нескольких регистрах содержится информация о видах деятельности и доходах. Наиболее важной информацией о видах деятельности населения, собираемой в ходе проведения переписей, являются данные о занятости (регистры трудовой деятельности) и образовании. В этих регистрах указывается связь между лицом и его/ее местом работы или учебы. В них содержится также дополнительная информация о видах деятельности лица, таких, как данные о "характере занятости" и "продолжительности рабочего дня", если это касается работающих лиц, и "виде учебы", если речь идет об учащихся. В регистрах лиц, касающихся безработных и пособий по социальному обеспечению, сообщается дополнительная информация о выполняемом виде деятельности. Кроме того, в налоговых регистрах содержится информация о различных видах доходов.

Регистры экономической деятельности

8. В отношении занятых лиц информация о рабочем месте, предприятии и отрасли экономики собирается на основе регистров экономической деятельности. В 1995 году в Норвегии был учрежден центральный регистр юридических лиц, в котором используются специальные идентификационные номера; в настоящее время такие номера используются и в других наиболее важных регистрах.

Улучшение сбора данных

9. В течение многих лет административные записи и регистры использовались в Норвегии в качестве источников социальной статистики. Однако в течение последних 10-15 лет объем имеющихся административных данных возрос. Основной причиной этого, безусловно, стало создание административных регистров в ряде новых областей. Кроме того, все большее число регистров содержат данные об отдельных лицах.

10. До недавнего времени большая часть социальной статистики в Норвегии могла быть получена из регистров или по крайней мере на основе объединения данных регистров, относящихся к одной и той же административной системе (например, налоговых

регистров). Возросший объем административных данных и возможность оперативной увязки данных, получаемых из различных систем, позволяют осуществлять комплексный сбор статистических данных, относящихся к различным областям.

11. В ходе переписи 1990 года СУН сочло необходимым провести специальное обследование для обеспечения надлежащего качества данных о рынке труда. В отношении переписи 2000 года принято решение в отношении всех тем полагаться на данные регистров. Принятие этого решения объясняется следующими причинами:

- улучшилось качество основных регистров;
- появились новые административные источники, к которым относятся вновь созданные регистры и уже существующие регистры, дополненные новыми переменными величинами;
- данные из различных регистров объединяются в рамках системы комплексных данных с целью улучшения качества ключевых переменных величин;
- СУН располагает большим объемом информации о качестве данных.

12. Вышесказанное отнюдь не означает, что качество всех соответствующих данных регистров в настоящее время является удовлетворительным. Поэтому улучшение качества данных регистров представляет собой одну из важнейших целей проведения переписи.

3. Улучшение качества данных

13. СУН на основе регистров собирает годовые статистические данные о большинстве тем, охватываемых переписью населения. Контроль за качеством осуществляется с помощью использования обычных систем сбора данных. Кроме того, в рамках проведения переписи предусматривается выделение дополнительных ресурсов для повышения качества данных отдельных регистров. Это позволяет включать в статистическую информацию новые переменные, данные о которых поступают из регистров.

Повышение качества данных каждого регистра

14. Качество данных каждого используемого административного источника, безусловно, имеет очень важное значение для обеспечения качества данных всей переписи населения. Все регистры содержат некоторую неточную информацию. Одним из основных видов ошибок являются проблемы, связанные с охватом переписью обследуемых единиц. На практике это означает отсутствие в регистре или неправильное включение в него обследуемых единиц. Кроме того, данные о некоторых переменных величинах в регистрах могут отсутствовать, быть неточными или вводить в заблуждение.

15. Основная задача СУН заключается в выявлении таких ошибок и информировании административных органов, отвечающих за ведение регистров. Безусловно, это один из важнейших видов деятельности в рамках обычной системы сбора данных СУН. В рамках проведения переписи используются некоторые дополнительные ресурсы для выявления и исправления ошибок, что позволяет улучшить качество. Процесс обнаружения ошибок упрощается, если осуществляется согласование данных из различных регистров.

Повышение качества посредством сочетания данных из нескольких регистров

16. Посредством сочетания данных из нескольких регистров мы можем лучше обеспечить повышение качества. Сочетание данных особенно полезно при решении проблем, связанных с охватом обследуемых единиц, но оно может также использоваться и для корректировки при отсутствии или неточности соответствующих данных. В качестве примера можно привести сочетание регистров данных о рынке труда и регистров различных видов доходов. При одновременном использовании информации из всех этих регистров можно улучшить качество переменных, характеризующих положение на рынке труда, а также переменных, характеризующих доходы.

Новые переменные, появляющиеся при сочетании регистров

17. В ходе переписи населения некоторые основные переменные могут быть получены непосредственно из административных источников (например, данные, касающиеся пола, возраста). Для обследования производных тем, таких, как "социально-экономические группы", необходимо использовать большое количество переменных из различных регистров. Однако даже в отношении других переменных, таких, как "выполняемый в настоящее время вид деятельности" ("занят на производстве", "является безработным", "не является частью рабочей силы"), необходимо сочетание данных из различных регистров.

18. В качестве примера мы будем использовать переменную, характеризующую "выполняемую в настоящее время деятельность". Основная проблема в данном случае состоит в том, как выявить занятых лиц. Какого-либо одного регистра, охватывающего все группы занятых лиц, не существует. К основным регистрам в данной области относятся:

1. регистр лиц наемного труда;
2. регистр безработных лиц;
3. налоговый регистр годовой заработной платы;
4. налоговый регистр доходов (лица, работающие не по найму).

19. Согласно определению МОТ, лица должны классифицироваться как занятые в том случае, если они работают по крайней мере один час в базисную неделю. Именно это определение используется при проведении обследований рабочей силы (ОРС). При переписи населения, а также при сборе годовых статистических данных о рынке труда на основе регистров преследуется цель составить такого рода классификацию на основе данных регистров. При составлении системы классификации занятости будет использоваться информация из всех вышеперечисленных регистров. Правила составления классификации являются довольно сложными. В настоящем документе описываются лишь основные принципы.

1) Единственный регистр, в котором содержится надежная информация о продолжительности выполнения той или иной работы, является регистр лиц наемного труда. Согласно этому регистру, лицо, имеющее работу в течение базисной недели, обычно классифицируется как занятое. Исключение составляют случаи, когда в налоговом регистре не содержится информации о заработной плате.

2) Информация о продолжительности выполнения той или иной работы, содержащаяся в налоговом регистре заработной платы, является менее надежной. В большинстве случаев сообщается о том, что такая продолжительность составляет целый год. Поэтому лица, которые имеют работу лишь в соответствии с этим регистром (а не регистром лиц наемного труда) в большинстве случаев должны классифицироваться либо как занятые, либо как безработные в течение целого года. В этой связи наиболее надежной является информация об общем размере заработной платы в течение года. Лица, размер заработной платы которых превышает определенный уровень, обычно классифицируются как занятые. Уровень заработной платы рассчитывается на основе информации ОРС.

3) В отношении самостоятельно занятых лиц информация о продолжительности такой работы вообще отсутствует (имеется лишь информация об уровне дохода за год). Классификация должна проводиться в соответствии с правилом 2; однако в данном случае вместо зарплаты используется переменная, характеризующая "доход самостоятельно занятого лица".

4) Лица, зарегистрированные безработными в течение базисной недели, обычно классифицируются как незанятые (подробности см. в главе 4).

20. Качество составленной с соблюдением указанных правил классификации анализируется на микроуровне с помощью данных ОРС, при этом корректируются и сами правила. Окончательный результат состоит в том, что 85% всех занятых лиц классифицируется на основе правила 1, 7% - правила 2 и 8% - правила 3.

Использование обследований для дополнения данных регистров

21. В некоторых случаях данные регистров нуждаются в дополнении посредством проведения обследований. Регистр образования не пополняется данными об образовании,

получаемом за рубежом. Поэтому в качестве составной части переписи 2000 года будет проведено обследование с целью сбора информации о лицах, получающих образование за рубежом. Аналогичное обследование было проведено в ходе переписи 1990 года.

22. Еще одним примером является использование информации ОРС при разработке переменных рынка труда. Как упоминалось выше, эти данные играют важную роль при классифицировании выполняемого в конкретный момент вида деятельности. Данные ОРС используются также для улучшения качества других переменных, касающихся рынка труда, таких, как "обычная продолжительность работы" и "вид деятельности".

4. Согласованность

23. При проведении переписи населения важно обеспечить согласованность переменных. Это имеет особенно важное значение при сборе данных переписи из различных административных источников, а не на основе заполнения вопросника. Под согласованностью данных регистров мы понимаем такое положение, при котором увязывание информации из различных регистров обеспечивает согласованную классификацию видов деятельности всех лиц.

24. Однако на практике использование такого общего определения является весьма сложным. Некоторые сочетания видов деятельности явно нелогичны и не могут быть приняты. Какое-либо лицо не может одновременно классифицироваться как занятое и безработное. Другие сочетания могут быть "законными", однако маловероятными. С другой стороны, некоторые сочетания видов деятельности являются весьма распространенными (многие студенты имеют работу). Основная задача состоит в том, чтобы определить, какие сочетания видов деятельности являются реальными и какие несоответствия существуют между данными регистров.

Виды несоответствия между данными регистров

25. Весьма полезно более внимательно изучить различные виды несоответствия, поскольку в их основе могут лежать различные причины.

Различные определения переменных

26. Несогласованность между данными регистров на практике зависит от определений сопоставляемых переменных. Необходимо сопоставлять официальные (или юридически признанные) определения различных переменных. Если согласно налоговому регистру какое-либо лицо получает зарплату в течение определенного периода времени, это означает, что данное лицо должно регистрироваться в качестве занятого. Однако критерии регистрации, используемые в этих двух административных системах, допускают ряд исключений из этого правила. Другими словами, не следует рассчитывать на согласованность данных о всех обследуемых лицах.

Различные совокупности

27. Несогласованность может возникать и при использовании различных совокупных данных. При проведении переписи населения эта проблема не стоит остро. Однако для регистров коммерческой деятельности, в которых единицами обследования выступают учреждения и предприятия, эти проблемы носят гораздо более серьезный характер.

Различия в процедурах отчетности

28. Официальные определения определяют характер содержания регистров. Реальное содержание зависит также от использования этих определений в административной системе. Это подчеркивает важность изучения процедур отчетности как составной части контроля качества.

Различные базисные периоды

29. Одной из серьезных причин несогласованности данных является использование в соответствующих регистрах различных базисных периодов. Собранные данные могут относиться к конкретному дню в одних регистрах (но необязательно к одному и тому же дню во всех регистрах), либо к более продолжительным периодам времени, например к одному году, в других регистрах.

Погрешности

30. Помимо вышеупомянутых случаев несогласованности данных, во всех регистрах может содержаться неточная информация.

Пример: несогласованность между данными регистра лиц наемного труда и регистра безработных

31. Возможности устранения несогласованности данных регистров будут показаны на примере увязки данных регистра лиц наемного труда и регистра безработных.

32. Если в обоих этих регистрах используется один и тот же персональный идентификационный номер, то увязывание содержащихся в них данных не составляет никаких проблем. Число лиц, имеющих недействительный идентификационный номер, крайне мало. Подлинная проблема состоит в получении согласованных данных о всех охваченных обследованием лицах. В случае совпадения указанных периодов согласование достигается посредством корректировки сроков либо трудовой деятельности, либо периода безработицы.

33. Согласно определению МОТ одно и то же лицо не может классифицироваться одновременно как безработное и как трудоустроенное. Однако мы обнаруживаем, что в отношении 1992 года 33% всех периодов безработицы совпадало со временем выполнения одной или нескольких работ.

34. Хотя это и противоречит административным правилам, некоторые лица могут быть зарегистрированными на бирже труда в качестве безработных даже в том случае, если в это время они выполняют какую-нибудь работу. Однако в большинстве случаев несогласованность данных возникает в результате ошибок в данных одного или обоих регистров. Это может объясняться двумя основными причинами:

1. Дата прекращения трудовой деятельности могла быть не сообщена для включения в регистр лиц наемного труда, или указанная дата прекращения или начала трудовой деятельности может быть неточной.
2. В регистре безработных лиц может указываться неточная дата прекращения срока, когда лицо было безработным.

35. На основе изучения используемой в этих двух регистрах практики отчетности мы разработали ряд правил для корректировки продолжительности указанных периодов. Для такой корректировки должны применяться следующие критерии:

- какой период начался первым (занятости или безработицы);
- продолжительность совпадения сроков одного и другого периодов;
- зарегистрировано ли прекращение трудовой деятельности.

36. При совпадении сроков в течение длительного периода обычно корректируется продолжительность трудовой деятельности, а при непродолжительном периоде совпадения сроков коррективы вносятся в данные о безработице. В тех случаях, когда о прекращении трудовой деятельности не сообщалось, коррективы вносятся в сроки ее выполнения (в данной ситуации нельзя рассчитать продолжительность времени совпадения обоих периодов).

37. При применении в целях корректировки такого приема число занятых в конкретный базисный день (в качестве даты было выбрано 1 декабря 1992 года) сокращается на 3,4%, а число безработных - на 2%. Предстоит провести оценку эффективности и использования данного приема; он еще не применялся для разработки статистических данных о рынке труда.

5. Измерение качества

38. Как указывалось выше, улучшение качества данных регистров является одной из основных целей проведения переписи. Не менее важное значение имеет также подготовка и представление информации о качестве получаемых данных.

Измерение качества имеет важное значение с двух точек зрения:

- с точки зрения контроля качества как составляющей части процесса подготовки статистических данных и
- с точки зрения описания качества публикуемых статистических данных.

39. Все статистические данные, получаемые на основе регистров и публикуемые в настоящее время СУН, сопровождаются своего рода качественным описанием. Цель переписи заключается в разработке стандартных способов описания качества всех переменных, используемых в ходе переписи.

40. Большинство критериев статистического измерения качества разработано для проведения обследований. Одной из важных составляющих переписи является разработка критериев измерения качества, в первую очередь подходящих для статистических данных, получаемых на основе регистров. При этом предполагается использовать данные методы применительно не только к статистическим данным переписи, но также и к другим статистическим данным, получаемым на основе регистров.

6. Обследование контроля качества

41. В целях измерения качества данных переписей необходимы дополнительные источники информации. За переписью населения традиционно проводится обследование вопросов, связанных с контролем качества. Его цель заключается в сборе данных о выборочных переменных с целью измерения качества соответствующих данных переписи. Такие данные могут собираться различным путем:

1. с помощью независимого выборочного обследования;
2. с помощью получения ответов на вопросы, дополняющие существующие обследования;
3. с помощью данных существующих обследований.

42. Традиционно обследования вопросов, связанных с контролем качества, проводятся в форме независимых выборочных обследований. План проведения переписи в 2000 году предусматривает использование второго и третьего указанных способов. Проведение переписи прежде всего на основе данных регистров, а не на основе заполнения вопросников позволяет более эффективно использовать данные проводимых обследований. В качестве примера мы можем использовать данные обследования рабочей силы (ОРС).

43. Ключевые переменные, характеризующие рынок труда, касаются выполнения вида деятельности в конкретный момент. Это означает, что мы собираем данные о положении на рынке труда каждого человека в течение базисной недели. Аналогичные данные

собираются и в рамках ОРС. Поэтому ОРС может использоваться в качестве источника данных для контроля качества используемых в ходе переписи переменных рынка труда. Число лиц, опрошенных в ходе ОРС в течение одной недели, не превышает 2 000 человек. Однако благодаря использованию регистров "данные переписи" могут быть получены не только в отношении недели проведения переписи (первая неделя ноября 2000 года), но также и в отношении всех других недель в течение более продолжительного периода времени, например последнего квартала 2000 года. Это дает нам возможность использовать результаты 24 000 опросов, проводимых в ходе ОРС.

44. В отношении переменных, описывающих "обычный вид деятельности" (применительно к положению в 2000 году), ОРС не может использоваться непосредственно как источник данных. В этой связи альтернативой будет включение в ОРС, которое планируется провести в первом квартале 2001 года, ряда дополнительных вопросов.

45. Мы считаем, что эти методы сбора данных сократят издержки проведения обследования вопросов, касающихся контроля качества, приблизительно на 30% по сравнению с проведением независимого обследования.

7. Выводы

46. Принятие решения получить все данные переписи 2000 года на основе регистров объясняется улучшением в течение последнего года качества данных регистров. Однако оно все еще остается неудовлетворительным. Поэтому одной из важнейших целей проведения переписи является уделение большего внимания аспектам качества. Это предполагает улучшение качества выборочных переменных, сочетание данных из нескольких регистров, обеспечение согласованности данных, полученных из различных источников, и стандартизацию измерения и описания качества. Если эти проекты обеспечения надлежащего качества принесут положительные результаты, то в следующем столетии СУН сможет ежегодно публиковать результаты переписи населения.
