

РАЗДЕЛ 1. ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ

ЭКОНОМЕТРИЧЕСКИЕ МОДЕЛИ УПОРЯДОЧЕННОГО ВЫБОРА – ПРИМЕР ОЦЕНОК УДОВЛЕТВОРЕННОСТИ ЖИЗНЬЮ

А.В. Аистов,

Л.А. Леонова

*Национальный исследовательский университет
«Высшая школа экономики» – Нижний Новгород*

Рассматривается применение моделей упорядоченного выбора для изучения экономических закономерностей. В качестве анализируемой дискретной упорядоченной зависимой переменной выбрано удовлетворение жизнью. Данная переменная, представляющая собой субъективные суждения человека о «счастье», пришла в экономику из психологии и нашла в ней свое место как прокси для функции полезности. Явное влияние ненаблюдаемых индивидуальных особенностей респондентов приводит к необходимости использования панельных данных и моделей с фиксированными эффектами.

Эмпирическая экономическая литература, посвященная самоотчету человека о счастье (удовлетворенности жизнью) развивалась крайне активно с 1970-х гг. Подробный анализ особенностей эконометрического тестирования взаимосвязей удовлетворенности жизнью представлен в статье [4]

Психологи придерживаются кардиналистской точки зрения в вопросе оценки уровня счастья: например, разница в счастье между 4 и 5 для индивидуума такая же, как между 8 и 9. В экономической литературе кардиналистская теория вызывает большое сомнение, поэтому стала устойчивой гипотеза о том, что ответы об уровне удовлетворенности лишь ординалистически сравнимы: то неизвестно, каково относительное различие между ответами об уровне удовлетворения, но все индивидуумы разделяют одинаковые интерпретации каждого возможного ответа.

Огромное практическое преимущество кардиналистского допущения в том, что каждый может легко выявить для рассмотрения

разницу в уровнях счастья и соотнести ее с изменением наблюдаемых переменных. Это значит, что любой эффект от стационарных (не зависящих от времени) ненаблюдаемых факторов выбывает в случае линейной спецификации. Как результат, существует много статей в психологической литературе, в которых вводится поправка на стационарные во времени ненаблюдаемые переменные, коррелированные с обозреваемыми, при этом используются эконометрические модели, учитывающие панельный характер данных с индивидуальными фиксированными эффектами.

При эконометрическом моделировании полезность полагается латентной зависимой переменной для степени удовлетворённости. В качестве объясняющих и контролирующих используется «традиционный» набор регрессоров: доходы, возраст, пол, уровень образования и т.п.

Если использовать в качестве примера данные Российского Мониторинга Экономического положения и Здоровья населения НИУ-ВШЭ (РМЭЗ), наблюдаемые значения латентной переменной формируются на основе ответа на вопрос: «Насколько Вы удовлетворены своей жизнью в целом в настоящее время?» В качестве ответа респонденту предлагается сделать выбор по 5-балльной шкале: полностью удовлетворены; скорее удовлетворены; и да, и нет; не очень удовлетворены; совсем не удовлетворены. В рамках общего анализа методологии обозначим ответ на этот вопрос через GS (general satisfaction).

Итак, пусть GS_{it} - эндогенная переменная для индивидуума i в момент времени t с уровнем счастья, снижающимся с ростом численного значения переменной (для РМЭЗ). Это дает возможность обозначить за 1 уровень «полностью удовлетворен», а все другие ответы за 0. Общий вопрос исследований счастья: установить, при определенных допущениях, влияние наблюдаемых характеристик x_{it} на GS_{it} при наличии ненаблюдаемых переменных ε_{it} .

Предположения.

1. Общие предположения.

Существуют три главные вариации допущений, используемых при интерпретации ответов об удовлетворении жизнью.

A1. Уровень удовлетворения – это положительная монотонная трансформация от подразумеваемой метафизической концепции, называемой благосостояние и обозначаемой $W(\cdot)$: если $GS_{it} > GS_{is}$, то $W_{it} > W_{is}$.

A2. Уровень удовлетворения порядково (ординалистически) сравним для различных индивидуумов: если $GS_i > GS_j$, то $W_i > W_j$.

А3. Уровень удовлетворения для различных индивидуумов сравним в кардиналистическом смысле: $(W_i - W_j) = \omega(GS_i - GS_j)$, где $\omega(\cdot)$ постоянный множитель (мультипликативная константа). Обычно $\omega(GS_i - GS_j)$ принимается как $(GS_i - GS_j)$.

Первое допущение предполагает соответствие между измеряемой переменной GS_{it} и метафизической концепцией, которая обычно интересует исследователей W_{it} . Если эмоциональное выражение и поведенческий выбор действительно соотносимы с подразумеваемой концепцией благосостояния, то GS может быть использована и как прокси для благосостояния.

Второе допущение, о порядковой сравнимости, подразумевает, что индивидуум разделяет общее мнение о том, что такое счастье. Оно основано на двух психологических заключениях. Во-первых, индивидуумы отчасти способны распознавать и предсказывать уровень удовлетворения других. Во-вторых, индивидуумы, принадлежащие к одной языковой группе, имеют общее понимание того, как переводить внутренние чувства в численную шкалу, проще говоря, выстраивать их по порядку, чтобы коммуницировать друг с другом. Эмпирический анализ GS при наличии ординалистической предпосылки, обуславливает применение моделей латентных переменных, таких как упорядоченные пробит и логит модели.

Третье допущение предотвращает экстремистские поведенческие ответы, в которых человек предполагается либо очень грустным, либо очень радостным, - в этом случае разница в благосостоянии между средними категориями крайне невелика. Когда GS предполагается кардиналистической мерой благосостояния, то эмпирический анализ часто реализуется средствами МНК.

2. Статистические предпосылки.

Статистические предпосылки имеют отправным пунктом существование и эффекты ненаблюдаемых факторов в наборе данных:

С1. Существуют изменяемые во времени ненаблюдаемые факторы, ε_{it} , соотносимые с наблюдаемыми переменными неизвестным способом.

С2. Существуют неизменные во времени ненаблюдаемые факторы v_i , относящиеся к внутреннему уровню наблюдаемых факторов, и существуют изменяющиеся со временем ненаблюдаемые факторы, ε_{it} , несвязанные с наблюдаемыми: $\text{cov}(\varepsilon_{it}, x_{it}) = \text{cov}(v_i, \Delta x_{it}) = 0$ и $\text{cov}(v_i, x_{it}) \neq 0$.

С3. Ненаблюдаемые факторы, ε_{it} и v_i , либо не связаны с наблюдаемыми факторами, либо их соотношение известно: $\text{cov}(\varepsilon_{it}, x_{it}) = z^1_{it}$ и $\text{cov}(v_i, x_{it}) = z^2_{it}$, где z - ноль или известная функция.

Первая предпосылка зачастую возникает вследствие экономической теории: индивидуумы постоянно принимают решения базируясь на принуждении и будущих ожиданиях. Нечто ненаблюдаемое, что влияет на GS и на изменения в ожиданиях, будет влиять на наблюдаемые решения.

Вторая предпосылка (C2) состоит в том, что все релевантные изменяющиеся во времени факторы подразумеваются наблюдаемыми. Остающиеся фиксированные ненаблюдаемые факторы влияют на уровень других переменных, хотя не на их изменения. Главные кандидаты на такие фиксированные ненаблюдаемые переменные в экономическом анализе – это индивидуальные черты индивидуума. В работах [5] и [1] приведены исчерпывающие психологические свидетельства о том, что постоянные черты – это лучшие предсказатели уровня удовлетворенности. В тоже время демографические и социоэкономические переменные в лучшем случае способны отловить лишь 15% разброса GS [4], гены и постоянные психологические черты как было найдено, имеют корреляцию с GS до 80% [16].

Третья предпосылка (C3) предполагает, что могут быть ненаблюдаемые факторы, но они либо ортогональны наблюдаемым переменным, либо относятся к тому, что наблюдаемо и известно и поэтому может быть проконтролировано. Данная предпосылка кажется разумной только в случае, когда используются рандомизированные данные.

3. Модели

Модели с предположением A3

Одна из популярных моделей, основанных на предположении A3 - это оценка:

$$GS_{it} = x_{it}\beta + \varepsilon_{it} \quad (1)$$

Здесь ε_{it} имеет нулевое математическое ожидание и ортогонален x_{it} , предполагается оценивание по МНК. В психологии очень популярной практикой является получение главного результата в виде таблицы корреляций между массивом GS_{it} и некоторыми наблюдаемыми характеристиками.

Эта модель была употреблена практически во всех психологических кросс-секционных исследованиях [1,5,22]. Среди экономистов данная спецификация является основной, лишь в случае сравнения совокупного удовлетворения между странами и, следовательно, по смыслу предполагается кардиналистичность [9,10,20,18,14,8].

Эта модель требует выполнения предпосылок А3 – С3. Преимущество использования А3 состоит в том, что достаточно легко ослабить С3 и перейти к С2 взятием и первой разницы (1):

$$GS_{it} - GS_{it-1} = \Delta x_{it} \beta + \Delta \varepsilon_{it} \quad (2)$$

Очевидно, что если ввести индивидуальную черту v_i , относящуюся к x_{it} , то она сократится в уравнение (2). Эта формулировка – стандартная модель причинности в психологической литературе, использующей панельные данные или временные ряды [6,1].

Из экономистов лишь Gelach and Stephan[12], Корп [15] ограничились в работах использованием подобной структуры МНК с фиксированными эффектами на уровне индивидуумов.

Модели с предположением А2

Основная модель при принятии предположения А2, то порядковой сравнимости, это модель с латентной переменной:

$$\begin{aligned} GS_{it} &= x_{it} \beta + \varepsilon_{it}, \\ GS_{it} &= k \Leftrightarrow \lambda_k \leq GS_{it} < \lambda_{k+1}, \\ \varepsilon_{it} &\perp x_{it} \end{aligned} \quad (3)$$

где GS^* – латентная переменная, а GS – наблюдаемый уровень удовлетворения. В зависимости от предполагаемого распределения ошибки ε_{it} данная формулировка приведет к упорядоченным логит или пробит моделям, которые могут быть оценены методом максимального правдоподобия. Для трактовки причинности должна также выполняться предпосылка С3. Эта модель наиболее часто используется в исследованиях экономистов.

В отличие от (1) эта модель не легко приспособляется легко к включению ненаблюдаемой индивидуальной неоднородности. Как показано Maddala [17] в упорядоченной пробит модели, наличие фиксированных индивидуальных эффектов приводит к несостоятельным оценкам. Фиксированные эффекты смещают оценки, и, следовательно, не существует оценки первой разницы для фиксированного эффекта в модели с латентной переменной. Это серьезно препятствует осуществлению предпосылки С2 для этой модели в контексте панельных данных. Существуют оценки условным методом максимального правдоподобия для логит модели с фиксированными эффектами, которые могут быть использованы, когда число различных категорий зависимой переменной уменьшается до двух.

Альтернатива предпосылке А2 – это принять конкретную структуру соотношения между не меняющимися во времени

ненаблюдаемыми и наблюдаемыми переменными. Первый вариант, предложенный Mundlak [19], специфицировать корреляцию между неменяющимися во времени ненаблюдаемыми переменными и меняющимися во времени наблюдаемыми переменными линейной функцией от наблюдаемых факторов. Van Praag [21] использовал: $\varepsilon_{it} = \alpha x_i + v_i + \eta_{it}$, где αx_i – улавливает корреляцию между фиксированными ненаблюдаемыми переменными и наблюдаемыми. Кроме того, можно применить упорядоченную логит или упорядоченную пробит модель со случайными индивидуальными эффектами, которые фиксированы во времени. В этом случае $\varepsilon_{it} = v_i + \eta_{it}$ с нормально распределенными v_i и η_{it} , ортогональными друг другу и также ортогональными наблюдаемым характеристикам x_{it} .

4. Порядковые модели с фиксированными эффектами.

Рассмотрим модели, для которых не требуется выполнения С3, для них достаточно выполнения С2. Такие А2 – С2 модели комбинируют нежелание экономистов принимать кардиналистичность с возможностью «кардиналистов» использовать оценки индивидуальных фиксированных эффектов. Этот подход базируется на работах [23] и [13].

$$GS_{it} = x_{it}\beta + f_i + \varepsilon_{it}, \quad (4)$$

$$GS_{it} = I(GS_{it} > 0).$$

Из чего следует бинарная модель с фиксированными эффектами. Для того чтобы воспользоваться данной моделью, необходимо преобразовать зависимую переменную так, чтобы она принимала только два значения. Так, Winkelmann and Winkelman сократили число уровней удовлетворенности с 11 (шкала от 0 до 10) до двух, приняв за ноль уровни ниже 7, а за единицу – остальные [23]. Авторы обеих работ, учитывая статистику [3], предложили следующую вероятность наблюдения GS_{i1}, \dots, GS_{iT} , условную по их сумме:

$$P[GS_{i1}, \dots, GS_{iT} | \sum_t GS_{it}, \beta, f_i, x_{it}] = \frac{\sum^{(GS_{it} x_{it})} \beta e^{-t}}{\sum_{GS \in S(\sum_t GS_{it})} \sum_{t=1}^T I(GS_{it} > k_i) x_{it} \beta e^{-t}}$$

Здесь $S(\sum_t GS_{it})$ обозначает набор всех возможных комбинаций

GS_{i1}, \dots, GS_{iT} , суммирование которых даёт $\sum_t GS_{it}$. Для $T=2$ это

означает вероятность равную $\frac{e^{(GS_{i1}x_{i1} + GS_{i2}x_{i2})\beta}}{e^{x_{i1}\beta} + e^{x_{i2}\beta}}$, причём

используются только индивидуумы для которых $GS_{i1} + GS_{i2} = 1$.

Так как эти модели используют только тех индивидуумов, кто пересек пороговое значение, установленное исследователем, существуют очень большие потери в данных. Опасность сокращения выборки в том, что стандартные ошибки коэффициентов растут и это приводит к большому разбросу остатков.

Другое ограничение обеих работ, в том, что они не включают дамми- переменные для времени, а значит, специфический фактор времени не контролируется.

Ferrer-i-Carbonell and Frijters обратились к этим ограничениям в статье [11] и путем расширения идеи Chamberlain [3] уделили внимание упорядоченной логит модели с фиксированными эффектами:

$$GS_{it} = x_{it}\beta + f_i + \varepsilon_{it},$$

$$GS_{it} = k \Leftrightarrow \lambda_k^i \leq GS_{it} < \lambda_{k+1}^i, \quad (5)$$

$$t = 1, \dots, T; k = 0, \dots, K; A(\varepsilon_{it}) = \frac{e^{\varepsilon_{it}}}{1 + e^{\varepsilon_{it}}} - c.d.f. \varepsilon_{it}.$$

Это упорядоченная логит модель с фиксированными индивидуальными эффектами и индивидуальными характеристическими порогами – λ_k^i . Все, что предполагается о значениях порогов, – это то, что они возрастают, то $\lambda_k^i < \lambda_{k+1}^i$ (в случае обратного упорядочивания альтернатив от лучшего к худшему, как в РМЭЗ, предполагается убывающий характер λ). Предположение о порядковой сравнимости выполняется. Это означает модель A1 – C2.

Статистика выглядит как

$$\begin{aligned}
& P[I(GS_{it} > k_i), \dots, I(GS_{iT} > k_i) / \sum_t I(GS_{it} > k_i) = c] = \\
& \frac{\prod_{t=1}^T \{1 + I(GS_{it} > k_i)\} [e^{-\lambda_{k_i}^i + (x_{it}\beta + f_i)} - 1]}{\prod_{t=1}^T [1 + e^{-\lambda_{k_i}^i + (x_{it}\beta + f_i)}]} = \\
& = \frac{\sum_{GS \in S(k_i, c)} \prod_{t=1}^T \{1 + I(GS_{it} > k_i)\} [e^{-\lambda_{k_i}^i + (x_{it}\beta + f_i)} - 1]}{\prod_{t=1}^T [1 + e^{-\lambda_{k_i}^i + (x_{it}\beta + f_i)}]} = \\
& = \frac{e^{\sum_{t=1}^T I(GS_{it} > k_i) x_{it}\beta}}{\sum_{GS \in S(k_i, c)} e^{\sum_{t=1}^T I(GS_{it} > k_i) x_{it}\beta}}
\end{aligned}$$

При этом $0 < c < T$ и $S(k_i, c_i)$ обозначает набор всех возможных комбинаций GS_{i1}, \dots, GS_{iT} для которых $\sum_t I(GS_{it} > k_i) = c_i$, где c_i отвечает за число временных периодов в которых уровень удовлетворения выше барьерного значения k_i .

Эти оценки аналогичны простой логит модели с фиксированными эффектами в том смысле, что данные сводятся к бинарной переменной, но затем применяются индивидуально в массиве данных через свободный параметр k_i . Это значит, что можно включить наблюдения всех индивидуумов, для которых оценка уровня удовлетворения меняется и, следовательно, дать гораздо более полный охват информации, используя K категорий взамен двух. Этот вариант получения оценок не может прогнозировать вероятность и предельный эффект без дополнительных допущений, например, что индивидуальный фиксированный эффект равен нулю.

В целом можно сделать заключение, что выбор в пользу предпосылки о кардиналистическом или ординалистическом характере ответов на вопрос об уровне удовлетворения незначительно влияет на результаты при оценке анализируемых факторов.

Эмпирический анализ на примере удовлетворенности жизнью

Для эконометрического анализа в качестве зависимой переменной использована мера удовлетворенности жизнью. Для этого использовался ответ на следующий вопрос: «Насколько Вы удовлетворены своей жизнью в целом в настоящее время?», на который предлагались варианты ответа:

- 1 полностью удовлетворены,
- 2 скорее удовлетворены,
- 3 и да, и нет,
- 4 не очень удовлетворены,
- 5 совсем не удовлетворены.

Базой данных анализа служил Российский Мониторинг Экономического положения и Здоровья населения НИУ-ВШЭ (RLMS-HSE) – панельный опрос, репрезентативный для России. Рассмотрен период с 1994 по 2009 г. Основная масса респондентов опрашивалась ежегодно в октябре-ноябре, в 1997 и 1999 гг. опрос не проводился.

Рассматриваемая модель имела вид

$$U_t = \beta_1 \ln(y_t) + Z_t' \delta + \varepsilon_t ,$$

где y_t – это реальный индивидуальный доход, а Z включает демографические и прочие характеристики, такие как возраст, возраст в квадрате, переменная урбанизации (проживание в городе или нет), образование, федеральный округ, состояние здоровья, статус занятости и пр. Удовлетворенность жизнью использована как прокси для полезности.

Упорядоченный выбор

Первым шагом в оценке модели являлось использование упорядоченной пробит модели по отдельным годам. В этом случае латентная переменная (прокси для полезности) имеет вид

$$y_i^* = x_i' \beta + \varepsilon_i ,$$

где x_i – вектор объясняющих переменных (включая статус занятости), β – вектор коэффициентов, ε_i – случайная компонента, распределенная по нормальному закону.

Наблюдаемая дискретная переменная определяется такой латентной непрерывной переменной, как

$$y_i = \begin{cases} 1, & \text{если } \tau_1 = -\infty \leq y_i * \tau_2, \\ 2, & \text{если } \tau_2 \leq y_i * \tau_3, \\ 3, & \text{если } \tau_3 \leq y_i * \tau_4, \\ 4, & \text{если } \tau_4 \leq y_i * \tau_5, \\ 5, & \text{если } \tau_5 \leq y_i * \tau_6 = \infty. \end{cases}$$

где $\tau_1, \tau_2, \dots, \tau_6$ - пороговые значения.

Тогда вероятность некого исхода может быть получена, как

$$\Pr(y_i = j | x_i) = F(\tau_{j+1} - x_i' \beta) - F(\tau_j - x_i' \beta).$$

Положим $y_{ij}=1$, если $y_i=j$, и будет равно $y_{ij}=0$ в противном случае. Тогда для выборки, включающей N наблюдений (y_i, x_i) , логарифм функции правдоподобия будет выглядеть как:

$$\ln L(\beta, \tau_2, \dots, \tau_5; y, x) = \sum_{i=1}^N y_{ij} \ln \Pr(y_i = j | x_i).$$

При этом, для упорядоченной пробит модели F – это кумулятивная функция стандартного нормального распределения.

Получены следующие результаты (табл.1).

Таблица 1

Оценки коэффициентов в ordered probit модели (Зависимая переменная: уровень удовлетворения жизнью)

Объясняющие переменные	1996	1998	2000	2004	2006	2008
Статус занятости	-	0.123** (0.057)	-0.037 (0.049)	0.153*** (0.043)	0.158*** (0.045)	0.242*** (0.039)
Логарифм доходов	-0.006 (0.006)	-0.087*** (0.014)	-0.040*** (0.008)	-0.027*** (0.008)	-0.142*** (0.017)	-0.037*** (0.007)
Пол (1- муж.. 0-жен.)	-0.149*** (0.033)	-0.166*** (0.044)	-0.074* (0.040)	-0.155*** (0.035)	-0.039 (0.034)	-0.117*** (0.031)
Возраст	0.037*** (0.005)	0.068*** (0.009)	0.058*** (0.008)	0.074*** (0.007)	0.077*** (0.007)	0.069*** (0.006)
(Возраст) ² /100	-0.034*** (0.005)	-0.068*** (0.011)	-0.063*** (0.010)	-0.075*** (0.009)	-0.081*** (0.008)	-0.072*** (0.007)
Проживание в городе	-0.113*** (0.037)	-0.116** (0.048)	-0.142*** (0.041)	0.064* (0.038)	0.111*** (0.035)	-0.032 (0.032)
Образование среднее общее	-0.029 (0.039)	0.040 (0.052)	0.005 (0.045)	-0.039 (0.041)	-0.012 (0.031)	-0.016 (0.029)

Окончание табл. 1

Объясняющие переменные	1996	1998	2000	2004	2006	2008
Среднее специальное образование	0.007 (0.036)	-0.068 (0.044)	-0.013 (0.039)	0.009 (0.0.35)	-0.044 (0.033)	0.004 (0.031)
Высшее образование	-0.073 (0.050)	-0.096 (0.062)	-0.071 (0.056)	-0.123** (0.048)	-0.111 (0.043)	-0.146*** (0.041)
Состояние в браке	-0.055 (0.039)	-0.054 (0.0.46)	-0.174*** (0.041)	-0.302*** (0.034)	-0.268*** (0.032)	-0.331*** (0.031)
Состояние здоровья (1 – были проблемы со здоровьем. 0- нет)	0.128*** (0.032)	0.077* (0.040)	0.059* (0.035)	0.243*** (0.0.3)	0.142*** (0.032)	0.143*** (0.028)
Москва и С.-Петербург	-0.145** (0.065)	-0.305*** (0.083)	-0.047 (0.086)	-0.316*** (0.064)	-0.135** (0.066)	-0.016 (0.059)
Центральный федеральный округ	-0.096* (0.058)	-0.218*** (0.072)	-0.069 (0.064)	-0.207*** (0.062)	-0.159*** (0.059)	-0.009 (0.055)
Южный ФО	-0.065 (0.063)	-0.144* (0.078)	0.054 (0.069)	-0.0.54 (0.064)	-0.125** (0.06)	0.129** (0.057)
Северо-западный ФО	-0.036 (0.066)	-0.221*** (0.080)	-0.052 (0.073)	-0.285*** (0.070)	-0.131** (0.066)	-0.284*** (0.062)
Дальневосточный ФО	-0.010 (0.081)	0.007 (0.102)	0.016 (0.087)	0.057 (0.085)	0.190** (0.082)	0.194** (0.078)
Сибирский ФО	0.074 (0.065)	0.130 (0.082)	0.256*** (0.070)	-0.033 (0.066)	-0.007 (0.065)	0.186*** (0.059)
Приволжский ФО	-0.030 (0.056)	-0.118* (0.069)	0.044 (0.062)	-0.087 (0.059)	-0.155*** (0.057)	-0.020 (0.052)
Органы управления	-0.446*** (0.096)	-0.534*** (0.110)	-0.480*** (0.088)	-0.481*** (0.086)	-0.315*** (0.081)	-0.490*** (0.074)
Наука/ образование	-0.099*** (0.029)	-0.106*** (0.036)	-0.165*** (0.033)	-0.112*** (0.029)	-0.105*** (0.028)	-0.127*** (0.026)
Промышленность	-0.078*** (0.019)	-0.051** (0.023)	-0.097*** (0.020)	-0.061*** (0.018)	-0.072*** (0.017)	-0.079*** (0.016)
Финансы	-0.028 (0.020)	-0.036 (0.023)	-0.060*** (0.021)	-0.068*** (0.018)	-0.042** (0.017)	-0.069*** (0.016)
Торговля	0.002 (0.014)	-0.034** (0.016)	-0.035** (0.014)	-0.024* (0.013)	-0.014 (0.011)	-0.034*** (0.011)
Сельское хозяйство	-0.080* (0.047)	-0.136** (0.056)	-0.080** (0.035)	-0.034 (0.040)	-0.023 (0.048)	-0.014 (0.033)
Строительство	-0.029*** (0.008)	0.001* (0.010)	-0.023*** (0.009)	-0.015* (0.008)	-0.013 (0.007)	-0.017** (0.007)
Транспорт	-0.021*** (0.007)	0.002 (0.008)	-0.039*** (0.007)	-0.013** (0.007)	-0.007 (0.006)	-0.019*** (0.006)
N	5464	3432	4179	5369	5942	6520
Prob>chi2	0.000	0.000	0.000	0.000	0	0
Pseudo R2	0.074	0.084	0.067	0.071	0.052	0.065

Примечание:

* - значимость коэффициентов на 10% уровне;

- ** - значимость коэффициентов на 5% уровне;
- *** - значимость коэффициентов на 1% уровне;

В скобках указано значение стандартной ошибки.

Панельный анализ

Персональные черты и другие индивидуальные ненаблюдаемые аспекты влияют на удовлетворенность. Индивидуальные эффекты, во-первых, позволяют учесть индивидуальные черты, частично устраняя эндогенность, вызванную тем, что черты характера индивида влияют не только непосредственно на уровень удовлетворения жизнью, но и на часть регрессоров. Во-вторых, позволяют учесть то, что пороговые значения у разных индивидов могут быть разными.

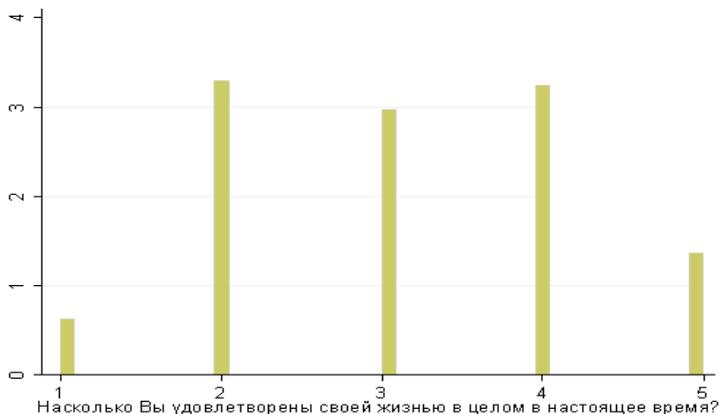
Так как индивидуальные черты стационарны (неизменны во времени), их может контролировать включение фиксированных эффектов.

Для дальнейшего эконометрического анализа была сформирована панель данных, включающая все доступные районы: с 5-го (1994 г.) по 18-й (2009 г.).

Модель бинарного выбора в панели

Как было сказано в анализе методологий, применяемых для оценки удовлетворенности жизнью, наибольшее внимание уделяется моделям с фиксированными эффектами. Пробит модели с фиксированными эффектами и приводят к несостоятельным оценкам параметров [2]. Поэтому была использована логит модель с фиксированными эффектами.

На первом этапе необходимо преобразовать зависимую переменную (удовлетворенность жизнью) из дискретной, имеющей пять градаций, в бинарную переменную. Для этого рассмотрим оценку распределения вероятности исходной переменной:



Из гистограммы можно предложить следующую схему формирования новой зависимой переменной:

$$binary_{satis} = \begin{cases} 0, & \text{если } satis = 1 \text{ или } satis = 2 \text{ или } satis = 3 \\ 1, & \text{если } satis = 4 \text{ или } satis = 5 \end{cases}$$

Таким образом, если полученная бинарная переменная принимает значение 1, то она отражает совсем неудовлетворенных и не очень удовлетворенных индивидуумов. В случае же, когда принимаемое значение равно 0, то соответствующие наблюдения отражают удовлетворенных людей.

После того как сформирована бинарная зависимая переменная, оценивались обычная процедура логит - оценивания, игнорирующая панельный характер данных, и затем логит регрессия с фиксированными эффектами.

Таблица 2

Оценки коэффициентов моделей бинарного выбора для панельных данных (Зависимая переменная: уровень удовлетворения жизнью)

Объясняющие переменные	Pooled logit	FE logit
Статус занятости	0,357*** (0,045)	0.250*** (0.078)
Логарифм доходов	-0,385*** (0,014)	-0.322*** (0.027)
Пол (1- муж., 0-жен.)	-0,314*** (0,024)	
Возраст	0,129*** (0,005)	-0.016 (0.020)

Продолжение табл.2

Объясняющие переменные	Pooled logit	FE logit
(Возраст) ² /100	-0,129 (0,006)	-0.151*** (0.023)
Проживание в городе	-0,032 (0,024)	
Образование среднее общее	-0,075 (0,051)	-0.266 (0.214)
Среднее специальное образование	-0,055** (0,026)	0.023 (0.197)
Высшее образование	-0,271*** (0,031)	0.092 (0.197)
Москва и С.-Петербург	-0,144*** (0,049)	
Центральный федеральный округ	-0.086* (0.044)	
Южный ФО	0.022 (0.046)	
Северо-западный ФО	0.121** (0.049)	
Дальневосточный ФО	0.433*** (0.048)	
Сибирский ФО	0.221*** (0.050)	
Приволжский ФО	0.040 (0.045)	
Органы управления	-0.694*** (0.061)	-0.288** (0.129)
Наука/ образование	-0.434*** (0.041)	-0.193* (0.106)
Промышленность	-0.091** (0.038)	-0.063 (0.084)
Финансы	-0.282*** (0.049)	0.039 (0.109)

Окончание табл.2

Торговля	-0.278*** (0.043)	-0.166* (0.100)
Объясняющие переменные	Pooled logit	FE logit
Сельское хозяйство	-0.515*** (0.160)	-0.597 (0.382)
Строительство	-0.117*** (0,040)	-0.131 (0.086)
Транспорт	-0.096 (0.124)	0.199 (0.282)
1998	1.443*** (0.049)	
2000	0.931*** (0.047)	
2001	0.317*** (0.046)	
2003	0.409*** (0.043)	
2004	0.184*** (0.044)	
2005	0.013 (0.044)	
2006	0.178*** (0.041)	
2007	0.078* (0.042)	
2008	0.003 (0.042)	
Константа	-1.247	
Число наблюдений	49 582	30 951
Число индивидуумов		5 525
Log L	-29146	-11 634

Примечание:

* - значимость коэффициентов на 10% уровне;

** - значимость коэффициентов на 5% уровне;

*** - значимость коэффициентов на 1% уровне.

В скобках указано значение стандартной ошибки.

Адаптированная пробит модель МНК

Логит модель с фиксированными эффектами, оцененная условным методом максимального правдоподобия, может быть основана только на подвыборках тех индивидуумов для которых есть изменение зависимой переменной во времени. Поэтому размер выборки обычно сокращается значительно, особенно в случае бинарной зависимой переменной. Учитывая это, условная логит модель отличается от линейной модели с фиксированными эффектами. В линейной модели с фиксированными эффектами наблюдения с неизменным значением зависимой переменной все-таки вносят вклад в оценку параметров наклона так долго, как они имеют некоторую вариацию в независимых переменных. Только наблюдения, которые не имеют изменений ни в независимых, ни в зависимой переменных, не способствуют оценке углов наклона в линейных моделях с фиксированными эффектами. Этот недостаток применения нелинейных моделей становится больше чем неудобство, когда неизменные во времени зависимые переменные удаляются из анализа вместе с индивидуальными фиксированными эффектами. В случае дискретной зависимой переменной, принимающей более двух значений, для устранения этой проблемы можно воспользоваться линейной моделью, где зависимая переменная «грубо кардинализована» с помощью методологии, которую предложили Van Praag/ Ferrer-i-Carbonel [21].

В принципе кардинализация порядковых переменных невозможна, но можно масштабировать переменную для того, чтобы сделать применение линейной модели более подходящим при определенных предпосылках. Метод probit adapted OLS (POLS), предложенный van в работе [21], состоит в извлечении квантилей стандартного нормального распределения, которые соотносятся с накопленными частотами различных категорий порядковой зависимой переменной.

Анализируемый уровень удовлетворенности жизнью имеет 5 градаций, со следующим распределением: $p(\text{satis}=1)=0.058$, $p(\text{satis}=2)=0.323$, $p(\text{satis}=3)=0.259$, $p(\text{satis}=4)=0.263$ и $p(\text{satis}=5)=0.097$.

Кумулятивные частоты в данном случае принимают значения $P(\text{satis}=1) = 0.058$, $P(\text{satis} \leq 2) = 0.381$, $P(\text{satis} \leq 3) = 0.64$, $P(\text{satis} \leq 4) = 0.903$, $P(\text{satis} \leq 5) = 1$, а соответствующие Z-values стандартного нормального распределения равны: $Z_{0.055} = -1.256$, $Z_{0.341} = 0,109$, $Z_{0.599} = 0.839$, $Z_{0.881} = 1.886$ и $Z_1 = \infty$.

Для данных значений оригинальной упорядоченной переменной, значения «кардинализованной» зависимой переменной конструируются взятием ожидания от стандартно распределенной нормальной переменной при условии, что она находится в интервале между двумя Z-values, корреспондирующими к классу значений оригинальной переменной. Тогда, присвоив индекс c кардинализованной переменной $satis$, обозначающей уровень удовлетворенности жизнью, получим

$$E(Z/Z < -1,256) = \frac{-\phi(-1,256)}{\Phi(-1,256)} \text{ если } satis = 1$$

$$E(Z/-1,60 < Z < 0,109) = \frac{[\phi(-1,256) - \phi(0,109)]}{[\Phi(0,109) - \Phi(-1,256)]} \text{ если } satis = 2$$

$$E(Z/0,109 < Z < 0,839) = \frac{[\phi(0,109) - \phi(0,839)]}{[\Phi(0,839) - \Phi(0,109)]} \text{ если } satis = 3$$

$$E(Z/0,839 < Z < 1,886) = \frac{[\phi(0,839) - \phi(1,886)]}{[\Phi(1,886) - \Phi(0,839)]} \text{ если } satis = 4$$

$$E(Z/1,886 < Z) = \frac{\phi(1,886)}{[1 - \Phi(1,886)]} \text{ если } satis = 5$$

где ϕ – функция плотности стандартного нормального распределения;
 Φ – кумулятивная функция стандартного нормального распределения.

Это приведет к:

$$satis_c = \begin{cases} -2.024, & \text{если } satis = 1 \\ -0.894, & \text{если } satis = 2 \\ -0.077, & \text{если } satis = 3 \\ 0.665, & \text{если } satis = 4 \\ 1.671, & \text{если } satis = 5 \end{cases}$$

В этом методе Z-values стандартного нормального распределения могут быть заменены порогами, взятыми из упорядоченной пробит или упорядоченной логит моделей. При этом используется информация модели в целом, а не только частоты распределения зависимой переменной, предназначенной для кардинализации.

После процесса кардинализации оценивается линейная модель с фиксированными эффектами. Таким образом, имеется возможность и

учесть наличие индивидуальных фиксированных эффектов, и избежать проблем, связанных с уменьшением выборки.

Таблица 3

Оценки на основе адаптированной пробит модели МНК(Зависимая переменная: уровень удовлетворения жизнью)

Объясняющие переменные	Pooled probit	FE
Статус занятости	0.193*** (0.021)	0.119*** (0.025)
Логарифм доходов	-0.219*** (0.006)	-0.100*** (0.008)
Пол (1- муж., 0-жен.)	-0.163*** (0.011)	
Возраст	0.075*** (0.002)	-0.018*** (0.006)
(Возраст) ² /100	-0.078*** (0.002)	-0.059*** (0.007)
Проживание в городе	-0.0002 (0.011)	
Образование среднее общее	-0.066*** (0.024)	-0.028 (0.067)
Среднее специальное образование	-0.027** (0.012)	0.022 (0.059)
Высшее образование	-0.118*** (0.014)	0.053 (0.057)
Москва и С.-Петербург	0.041* (0.022)	
Центральный федеральный округ	0.050** (0.020)	
Южный ФО	0.110*** (0.021)	
Северо-западный ФО	0.107*** (0.023)	
Дальневосточный ФО	0.303*** (0.023)	

Продолжение табл.3

Объясняющие переменные	Pooled probit	FE
Сибирский ФО	0.209*** (0.024)	
Приволжский ФО	0.143*** (0.021)	
Органы управления	-0.407*** (0.027)	-0.132*** (0.038)
Наука/ образование	-0.240*** (0.019)	-0.123*** (0.033)
Промышленность	-0.052*** (0.018)	-0.013 (0.027)
Финансы	-0.174*** (0.024)	-0.020 (0.034)
Торговля	-0.140*** (0.020)	-0.057* (0.031)
Сельское хозяйство	-0.274*** (0.076)	-0.221** (0.112)
Строительство	-0.057*** (0.019)	-0.049* (0.028)
Транспорт	-0.081 (0.057)	0.017 (0.085)
1998	0.871*** (0.023)	
2000	0.519*** (0.023)	
2001	0.153*** (0.022)	
2003	0.189*** (0.020)	
2004	0.102*** (0.020)	
2005	0.023 (0.020)	

Окончание табл.3

Объясняющие переменные		Pooled oprobit	FE
2006		0.089*** (0.019)	
2007		0.057*** (0.019)	
2008		0.013 (0.019)	
Число наблюдений		49 582	49 582
Число индивидуумов			13 505
Log L		-68 104	
Pseudo R ²		0.05	
Prob> χ^2		0.000	
R ²	within		0.075
	between		0.019
	overall		0.003
Prob>F			0.000

Примечание:

- * – значимость коэффициентов на 10% уровне;
- ** – значимость коэффициентов на 5% уровне;
- *** – значимость коэффициентов на 1% уровне;

В скобках указано значение стандартной ошибки.

В работе использованы результаты, полученные в ходе выполнения проекта №11-04-0055 «Занятость, доверие, удовлетворенность – поиск закономерностей», выполненного в рамках программы «Научный фонд ГУ-ВШЭ» 2011-2012 гг.

Список литературы

1. *Argyle M.* Causes and correlates of happiness, in Kahneman et al. 1999, chapter 18.
2. *Baltagi Badi H.* A companion to theoretical econometrics, Companions to Contemporary Economics. Malden, Mass. and Oxford: Blackwell, 2001. P, 709.

3. *Chamberlain G.* Analysis of covariance with qualitative data, *Review of Economic Studies*. 1980. vol. 47. P. 225–238.
4. *Diener E.* Subjective well-being // *Psychological Bulletin*. 1984. vol. 95. P. 542–575.
5. *Diener E. and Lucas R.E.* Personality and subjective well-being, in Kahneman et al. 1999.
6. *Diener E. and Suh E.* National differences in well-being, in Kahneman et al. 1999, chapter 22.
7. *Diener E., Suh E., Lucas R. and Smith H.* Subjective well-being: three decades of progress // *Psychological Bulletin*. 1999. vol. 125. P. 276–302.
8. *Di Tella R., MacCulloch R.J. and Oswald A.J.* Preferences over inflation and unemployment: evidence from surveys of happiness // *American Economic Review*. 2001. vol. 91. P. 335–341.
9. Easterlin R. Does economic growth improve the human lot? Some empirical evidence, in (P. David and R. Reder, eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press. 1974.
10. *Easterlin R.* Will raising the incomes of all increase the happiness of all? // *Journal of Economic Behavior and Organization*. 1995 vol. 27. P. 35–47.
11. *Ferrer-i-Carbonell A., Frijters P.* How important is methodology for the estimates of the determinants of happiness? // *The Economic Journal*. 2004. No. 114. P. 641–659.
12. *Gerlach K. and Stephan G.* A paper on unhappiness and unemployment in Germany *Economics Letters*. 1996. vol. 52. P. 325–330.
13. *Hamermesh D.S.* The changing distribution of job satisfaction // *Journal of Human Resources*. 2001. vol. 36. P. 1–30.
14. *Kenny C.* Does growth cause happiness, or does happiness cause growth? *Kyklos*. 1999. vol. 52. P. 3–25.
15. *Korpi T.* Is utility related to employment status? Employment, unemployment, labor market policies and subjective well-being among Swedish youth // *Labour Economics*. 1997. vol. 4. P. 125–147.
16. *Lykken D. and Tellegen A.* Happiness is a stochastic phenomenon // *Psychological Science*. 1996. vol. 7. P. 186–189.
17. *Maddala G.S.* *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.
18. *Micklewright J. and Stewart K.* Is the well-being of children converging in the European Union? // *Economic Journal*. 1999. vol. 109. P. F692–714.

19. *Mundlak Y.* On the pooling of time series and cross section data // *Econometrica*. 1978. vol. 46. P. 69–85.
20. *Oswald A.J.* Happiness and economic performance *Economic Journal*. 1997. vol. 107. P. 1815–1831.
21. *Van Praag B.M.S., Frijters P. and Ferrer-i-Carbonell A.* The anatomy of subjective well-being // *Journal of Economic Behavior and Organization*. 2001. vol. 51. P. 29–49.
22. *Veenhoven R.* Quality-of-life in individualistic society: a comparison of 43 nations in the early 1990's // *Social Indicators Research*. 1997. vol. 48. P. 157–186.
23. *Winkelmann L. and Winkelmann R.* Why are the unemployed so unhappy? Evidence from panel data // *Economica*. 1998. vol. 65. P. 1–15.

ЭКОНОМЕТРИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ ВЛИЯНИЯ ПАРАМЕТРОВ ВНЕШНЕЙ СРЕДЫ НА КАЧЕСТВО ИНСТИТУТОВ ГЕНДЕРНОЙ ВЛАСТИ

Е.В. Базуева

*Пермский государственный национальный
исследовательский университет, Пермь*

В современной экономике, институциональная среда гендерной власти¹ создает институциональные возможности выступать в качестве субъектов гендерной власти следующим экономическим агентам: мужчинам, социальному окружению, организациям и государству [1,2,3]. Каждый из них, формируя собственную систему ограничений для объекта власти в отведенной области регулирования властных отношений, выступает в качестве института гендерной власти².

¹ Под гендерной властью, на наш взгляд, следует понимать потенциальную возможность субъекта власти воздействовать на поведение объекта власти посредством угрозы применения санкций с целью максимизации собственной функции полезности (присвоения ренты власти) в условиях асимметрии распределения экономических ресурсов [См. об этом подробнее [1, 2].

² Под институтом гендерной власти мы понимаем особый вид социально-экономического института, сформировавшегося под влиянием объективных условий материального производства (общественного разделения труда), который можно представить в виде системы относительно устойчивых властных отношений, закрепленных через систему законодательных актов, контрактов и неформальных