Регрессия с эффектами взаимодействия:
детерминанты выбора STEM-дисциплин школьницами

Едва ли можно оспорить, что в современном мире тема профессиональной ориентации человека становится одной из ключевых. С одной стороны, правильный профессиональный выбор может стать основным фактором в построении успешного будущего конкретно взятого человека. С другой стороны, выбор профессии предопределяется и внешними факторами: развитием отраслей экономики так же, как нормами, ценностями и ритуалами вхождения в профессию. Однако современное общество до сих пор не предлагает равные возможности для мужчин и женщин: многие годы наблюдается явление поляризации научных карьер по гендерному признаку. Особенно остро эта тенденция проявляется в приоритетной и высокооплачиваемой STEM-сфере, вбирающей в себя естественные науки, технологии, инженерию и математику. Официальная статистика за 2016 год подтверждает превалирование мужчин, указывая на данные о 24% женщин, занятых в STEM в России и в Европе. Опыт зарубежных исследований указывает, что истоки данной проблемы берут свое начало ещё в школьные годы.[[1]](#footnote-1)

Проведенный опрос 700 школьников 5-х, 8-х и 11-х классов г. Москвы и г. Губкина (Белгородская область) подтвердил наличие гендерной асимметрии в предварительном выборе STEM-дисциплин учащихся общеобразовательных школ. Так, полученные данные указали, что приоритетными направлениями для мальчиков оказались естественные и технические науки, отвечающие за STEM-дисциплины. Девочки своим выбором продемонстрировали противоположную тенденцию, отдав предпочтение гуманитарным наукам и творческим направлениям. Только 35% школьниц остановили свой выбор на STEM-дисциплинах, в то время как мальчики превысили этот показатель практически вдвое (65%).

Для установления детерминант выбора STEM-дисциплин школьницами на собранных данных была построена модель бинарной логистической регрессии. Однако модель, включающая в себя только главные эффекты (то есть предикторы в их исходном виде, без учета сочетаний их категорий) обладала низкой прогностической силой: значение аналога коэффициента детерминации (псевдо-R2 Нагелькерке) составило 0,14. В рамках исследования, осуществляемого научно-учебной группой, было принято решение построить более точную модель регрессии, обратившись к эффектам взаимодействия. Это позволило увеличить прогностическую способность модели до значения псевдо-R2 0,51 без обращения к дополнительным предикторам.

*Алгоритм построения регрессионной модели
c эффектами взаимодействия*

Одной из гипотетических причин низкой объяснительной способности регрессионных моделей может выступать нелинейная форма связи между игреком (зависимой переменной) и иксами (предикторами). (Следует учесть, что логистическая регрессия, тоже предполагает линейную связь между зависимой (логитом) и независимыми переменными). Для того чтобы добиться нелинейности связи необходимо либо подобрать правильную нелинейную функцию, либо обратиться к эффектам взаимодействия, которым посвящена дальнейшая работа.

Первым шагом для поиска значимых эффектов взаимодействия было обращение к деревьям классификации. Один из наиболее подходящих для этого методов – ChAID. Он позволяет найти самые сильные взаимодействия предикторов, способных объяснить зависимую переменную – предварительный профессиональный выбор (факт выбора или не выбора ученицами траектории, связанной со STEM). Поскольку в основе ChAID лежит критерий хи-квадрат, независимые переменные могут быть измерены в любом типе шкал. Деревья классификации указали на 13 значимых связей между предикторами, с которыми необходимо работать дальше.

Поскольку деревья классификации ориентированы лишь на поиск самих взаимодействий, но не содержат в себе расчет параметров модели и «строгих» критериев качества предсказаний, следующим шагом стало обращение к многомерному дисперсионному анализу – методу, который позволяет сориентироваться в максимально возможной прогностической способности конкретного набора предикторов независимо от формы связи между ними и зависимой переменной. Так, посредством дисперсионного анализа были отобраны из 13 эффектов взаимодействия и 16 главных эффектов значимые, обладающие самой высокой объяснительной способностью. Данная процедура проходила путем пошагового удаления предикторов, имеющих самый низкий средний квадрат (mean square) – показатель, который можно рассматривать как нормированную на количество категорий величину силы эффекта.

Важное техническое преимущество работы с дисперсионным анализом состоит в том, что этот метод, в отличие от регрессии, не требует наличия в базе переменных, «отвечающих» за те или иные взаимодействия (они создаются виртуально). Метод позволяет отобрать ограниченный набор необходимых эффектов взаимодействия, и только они затем требуют подготовки в виде создания дополнительных переменных. После их соответствующей подготовки и перед запуском бинарной логистической регрессии все полученные сочетания были проверены на наполненность, и из дальнейшего анализа были исключены те из них, которые представлены менее, чем 15 респондентами. Это объясняется тем, что включение таких «редких» сочетаний в модель неизбежно привело бы к ее неустойчивости (невоспроизводимости на других данных).

Весь описанный алгоритм был направлен на отбор обладающих наибольшей объяснительной способностью эффектов взаимодействия, позволяющих построить высокоточную бинарную логистическую регрессию и углубить содержательный анализ. Данные шаги позволили увеличить предсказательную способность модели более чем в три раза (с 0,14 до 0,51) без обращения к дополнительным предикторам: иными словами, было установлено, что тот набор переменных, который исходно не показывал высокой прогностической силы, в действительности же ею обладал, но **только при условии учета взаимодействий этих переменных**.

*Интерпретация интересных эффектов*

*Девочки* **не** *выбирают STEM*, если они обучаются в классе с социально-экономическим профилем, при этом они не получают помощи в изучении математики от своих родителей, при этом в их школе выражена гендерно-ориентированная организация учебной жизни[[2]](#footnote-2).

*Девочки выбирают STEM,* если они обучаются в классе без профиля, оценивают свои знания в математике на самом высоком уровне *(*«я схватываю новые знания в математике лучше, чем мои одноклассники»), при этом они получают помощь от родителей в выполнении домашнего задания по математике, при этом они считают математику «женской» дисциплиной и при этом, по их оценке, гендерно-направленная организация учебной жизни в школе слабо выражена.

Важно отметить, что полученные эффекты взаимодействия являются нелинейными. Это значит, например, что гендерно-направленная организация учебной жизни сама по себе не оказывает достаточного влияния на зависимую переменную для высокой объяснительной способности модели. Однако в сочетании с двумя другими категориями переменных – профиль класса и получение помощи от родителей – влияние гендерно-направленной организации учебной жизни на выбор STEM-дисциплин достаточное для получения высокого псевдо-R2.

**Приложение**

Литература, использованная в ходе изучения гендерной асимметрии в предварительном выборе STEM-дисциплин школьниками:

1. *Ambrose R., Levi L., Fennema E.* The complexity of teaching for gender equity // Multicultural and gender equity in the mathematics classroom: The gift of diversity. 1997. P. 236-242.
2. *Archer L, DeWitt J, Osborne J, Dillon J, Willis B, Wong B*. Science aspirations, capital, and family habitus: how families shape children’s engagement and identification with science *//* American Educational Research Journal. 2012. Vol. 49. No. 5. P. 881–908.
3. *Aschbacher P, Li E, Roth E*. Is science me? High school students’ identities, participation and aspirations in science, engineering, and medicine // Journal of Research in Science Teaching. 2010. Vol.47. No. 5. P. 564–82.
4. *Bandura A., et al.* Self-efficacy beliefs as shapers of children’s aspirations and career trajectories // Child Development. 2001. Vol 72. P. 187–206.
5. *Buffington C., Cerf B., Jones C., Weinberg B.* STEM Training and Early Career Outcomes  of Female and Male Graduate Students: Evidence from UMETRICS Data Linked to the 2010 Census // American Economic Review. 2016. Vol. 106, No. 5. P. 333-338.
6. *Carlone H., Johnson A.* Understanding the science experiences of successful women of color: science identity as an analytic lens // Journal of research in science teaching. 2007. Vol. 44. No. 8. P.1187–1218.
7. *Catsambis S*. The Path to Math: Gender and Racial-Ethnic Differences in Mathematics Participation from Middle School to High School //Sociology of Education. 1994. Vol. 67. No. 3. P. 199-215.
8. *Ceci S., Williams W.* Why Aren’t More Women in Science? // Washington DC: American Psychological Association. P. 199–210.
9. *Coleman J*. Social capital in the creation of human capital // American Journal of Sociology*.* 1988. Vol. 94. P. 95–120.
10. *Correll S*. Gender and the Career Choice Process: The Role of Biased Self-Assessments // American Journal of Sociology. 2001. Vol. 106. No. 6. P. 1691-1730.
11. *Dee T.* How a Teacher's Gender Affects Boys and Girls. // Education Next, Fall. 2006. P. 69- 75.
12. *Eccles J.* Expectancies, values, and academic behaviors // San Francisco, CA: Achievement and achievement motives: Psychological and sociological approaches. 1983. P. 75-146.
13. *Eccles J., Vida M.* Predicting Gender and Individual Differences in College Major, Career Aspirations, and Career Choice // Society for Research on Child Development. 2003.
14. *Fennema E., Carpenter T., Jacobs V., Franke M., Levi L*. A longitudinal study of gender differences in young children's mathematical thinking // Educational Researcher. 1998. Vol.27. No. 5. P. 6-11.
15. *Fennema E., Peterson P*. Teacher - student interactions and sex-related differences in learning mathematics // Teaching and Teacher Education. 1986. Vol.2. No. 1. P. 19-42.
16. *Forgasz H., Leder G., Kloosterman P.* New Perspectives on the Gender Stereotyping of Mathematics // Mathematical Thinking and Learning. 2004. Vol.6. No. 5. P. 389-420.
17. *Gabay-Egozi L., Shavit Y., Yaish M*. Gender Differences in Fields of Study: The Role of Significant Others and Rational Choice Motivations // European Sociological Review. 2015. Vol. 31. No. 3. P. 284–297.
18. *Gonzalez H., Kuenzi J.* Science, technology, engineering, and mathematics (STEM) education: a primer. // Congressional Research Service // URL: http://www.fas.org/sgp/crs/misc/R42642.pdf. Проверено 26.11.2016.
19. *Gottfredson L., Lapan R.* Assessing gender based circumscription of occupational aspirations // Journal of Career Assessment. 1997. Vol. 5. P. 419–441.
20. *Graham M., Byars-Winston A., Hunter A., Handelsman J*. Increasing persistence of college students in STEM // Science. 2013.Vol. 341 P. 1455–1456.
21. *Herrnstein R., Murray C.* The Bell Curve: Intelligence and Class Structure in American Life // New York: Free Press. 1994.
22. *Hyde J., Fennema E., Lamon S.* Gender differences in mathematics performance // Psychological Bulletin. 1990. P. 139-155.
23. *Legewie J., DiPrete T.* High School Environments, STEM Orientations, and The Gender Gap in Science And Engineering Degrees // New-York: Columbia University. 2012.
24. *Legewie J., T. DiPrete*. The High School Environment and the Gender Gap in Science and Engineering // Sociology of Education. 2014. Vol. 87. No. 4. P. 259-280.
25. *Lynn R.* Racial and ethnic differences in intelligence in the United States on the Differential Ability Scale // Personal. Individ. Differ. 1996. Vol. 20. No. 2. P.217–273.
26. *Maccoby E., Jacklin C.* The psychology of sex differences // Stanford: Stanford University Press. 1974.
27. *McDaniel A*. The Role of Cultural Contexts in Explaining Cross-National Gender Gaps in STEM Expectations // European Sociological Review. 2016. Vol. 32. No. 1. P 122–133.
28. *Morgan S.* On the Edge of Commitment: Educational Attainment and Race in the United States // Stanford: Stanford University Press. 2005.
29. *Noddings N.* Perspectives from feminist philosophy // Educational Researcher. 1998. Vol.27. No. 5. P. 17-18.
30. OECD PISA Database // URL: http://www.oecd.org/pisa/pisaproducts/48852548.pdf. Проверено 25.11.2016.
31. OECD. The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence, PISA, OECD Publishing (2015) // URL: http://dx.doi.org/10.1787/9789264229945-en. Проверено 25.11.2016.
32. *Seligman L., Weinstock L*. The career development of 10 year olds // Elementary School Guidance and Counseling. 1991. Vol. 25. P. 172–182.
33. *Sewell W., Haller A., Portes A.* The educational and early occupational attainment process // American Sociological Review. 1969. Vol. 34. P. 82–92.
34. *Seymour E, Hewitt N.* Talking about Leaving: Why Undergraduates Leave the Sciences // Boulder, CO: Westview. 1997.
35. *Shelley J.* Gender and the Career Choice Process: The Role of Biased Self-Assessments // American Journal of Sociology. 2001. Vol. 106, No. 6. P.1691-1730.

*Smetackova I*. Gender Stereotypes, Performance and Identification with Math // Procedia - Social and Behavioral Sciences. 2015. Vol. 190. P. 211 – 219.

1. *Stanley J., Benbow* C. Sex differences in mathematical ability: Fact or artifact // Science. 1980. Vol. 210. No. 4475. P. 1262-1264.
2. *Tinto V.* Leaving College: Rethinking the Causes and Cures of Student Attrition // Chicago: Univ. Chicago Press. 1987.
3. *Xie Y., Fang M., Shauman K*. STEM Education // Annual Review of Sociology. 2015. No. 41. P. 331–357.
4. *Бем С*. Линзы гендера: Трансформация взглядов на проблему неравенства полов / Пер. с англ. М.: Российская политическая энциклопедия (РОССПЭН), 2004. С. 189.
5. *Дрозд В.* Гендерные различия в усвоении математики: реальность или иллюзия? // Didactics of mathematics: Problems and Investigations. 2008. Issue # 29. P.124-134.
6. *Здравомыслова Е., Темкина А* Социальное конструирование гендера как феминистская теория // Социологический журнал. 1998. № 3-4. С.171-182.
7. *Здравомыслова Е., Темкина А.* Социальная конструкция гендера и гендерная система в России // Материалы Первой российской летней школы по женским и гендерным исследованиям «Валдай 96». М., 1997. С.84-85.
8. *Киммел М.* Гендерное общество / Пер. с англ. М.: Российская политическая энциклопедия (РОССПЭН), 2006. С. 464.

*Малышева М*. Естественные и технические науки для женщин в XXI веке // Народонаселение. 2016. № 3. С. 76-85.

1. *Савинская О.* Гендерное равенство в STEM-программах дошкольного образования как фактор успешного технологического развития России // Женщина в российском обществе. 2016. №3. С. 16-24.

Федеральная служба государственной статистики. Сборник «Женщины и мужчины России». 2014 г. // URL: http://www.gks.ru/bgd/regl/b14\_50/Main.htm. Проверено 20.11.2016.

1. *Хасбулатова О.* Гендерные аспекты развития STEM-образования в России // Женщина в российском обществе. 2016. №3. С. 3-15.
2. *Шумилова О*. Политическое поведение женщин в условиях реформирования современного российского общества: Монография / Под общей ред. д-ра полит. наук, проф. И.А. Батаниной. Тула: ТулГУ, 2005. С. 182.
1. *Catsambis S*. The Path to Math: Gender and Racial-Ethnic Differences in Mathematics Participation from Middle School to High School //Sociology of Education. 1994. Vol. 67. No. 3. P. 199-215. *Correll S*. Gender and the Career Choice Process: The Role of Biased Self-Assessments // American Journal of Sociology. 2001. Vol. 106. No. 6. P. 1691-1730. *Eccles J., Vida M.* Predicting Gender and Individual Differences in College Major, Career Aspirations, and Career Choice // Society for Research on Child Development. 2003. [↑](#footnote-ref-1)
2. Переменная была получена с помощью метода главных компонент на предыдущих этапах анализа. В этом факторе основной акцент делается на содержании школьных уроков. К примеру, сюда входили уроки труда, где мальчикам и девочкам предлагается выполнение заданий, напрямую отражающих их гендерные различия. [↑](#footnote-ref-2)