НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет экономических наук

**Отчёт по исследовательскому проектному семинару**

**«Избранные темы семейной и гендерной экономики»**

Выполнила:

Алексеева Екатерина Игоревна

Руководитель:

Лукьянова Анна Львовна

г. Москва -2023

Оглавление:

[Введение 3](#_Toc136796959)

[Обзор журнала 3](#_Toc136796960)

[Анализ журнала за выбранный период с 2015 по 2017 год 6](#_Toc136796961)

[Выводы 17](#_Toc136796962)

# **Введение**

В течение второго семестра ИПС «Гендерные разрывы в сфере научной деятельности» был проанализирован авторский состав журнала «Социологические исследования», а также выполнялись задания по подсчётам различных показателей с целью выявления гендерного дисбаланса в публикации и цитируемости статей. Использовались такие инструменты, как Python, в котором делались все основные расчеты, и Excel, где строились графики по полученным значениям. В Python использовались такие библиотеки, как Pandas, Numpy, json и другие. Все подробные расчеты показателей представлены в отдельном файле Jupyter Notebook.

# **Обзор журнала**

Данный журнал был создан в 1974 году. В нём публикуются статьи о предмете социологии, её понятиях и инновациях в науке. Цель журнала, как заявлено на сайте, – познакомить читателей с новыми разработками социологии, идеями, проектами и техниками. Материал издания направлен на исследователей, научных работников, преподавателей, изучающих на теоретическом и эмпирическом уровнях социологические проблемы, в том числе современного общества.

Новый выпуск выходит каждый месяц. Журнал обладает таким ISSN: 01321625. Также «СоцИс» индексируется в РИНЦ, SCOPUS и Web Of Science.

Журнал стал индексироваться в SCOPUS с 2003 года. С момента присвоения номера до 2016 года «Социологические исследования» находились в 4 квартиле по качеству статей, в 2017 году – он перешел в 3-ий, а уже в следующем году - во второй квартиль. С 2018 «СоцИс» не менял больше своего положения. Что касается рейтинга SJR, показывающего авторитетность журнала и цитируемость, то в таблице 1 приведены его наилучшие значения по квартилям. Видно, что максимального значения этот показатель достиг в 2020 году, когда журнал находился во втором квартиле. Значения этого рейтинга меньше 1 означает пониженный относительно среднего потенциал журнала, однако в последние годы видна положительная динамика для этого показателя (см. Рис. 1).

Таблица 1 – Наилучшие значения рейтинга SJR из квартилей журнала

|  |  |  |
| --- | --- | --- |
| Квартиль | Год | Наилучшее значение SJR |
| 4 | 2016 | 0,118 |
| 3 | 2017 | 0,223 |
| 2 | 2020 | 0,356 |

Изображение выглядит как текст, снимок экрана, линия, График

Автоматически созданное описание

Рис. 1. Динамика рейтинга SJR для журнала «Социологические исследования»

Сотрудничество с зарубежными авторами является также важным показателем качества журнала. За весь период индексации в SCOPUS видна положительная динамика в проценте статей, написанных с участием зарубежных авторов, но увеличивалась она не планомерно. Интересно, что наибольшая доля статей, написанная с участием иностранных авторов, также была в 2020 году (см. Рис. 2).

Изображение выглядит как текст, снимок экрана, Шрифт, линия

Автоматически созданное описание

Рис. 2. Процентное изменение статей, написанных с участием иностранных авторов

Состав журнала также представлен на сайте: в редакционную коллегию входит 33 члена, из которых 15 – женщины. То есть доля женщин в редакции составляет 0.45, что близко к середине и говорит при первичном анализе о более-менее равном положении женщин и мужчин в составе журнала. Стоит заметить, что главный редактор и его заместитель – это мужчины, что также может влиять на политику отбора и публикации статей.

**Создание базы данных**

При помощи заданного кода в Python и парсинга через API сайта журнала в OpenAlex была получена информация о статьях журнала: год публикации, авторы, doi и другие характеристики. Затем была составлена таблица из необходимых признаков за период с 2015 по 2017 годы, в частности больше всего внимания было уделено именам авторов: с помощью цикла и анализа окончаний фамилий и отчеств авторов был определен их пол. После чего, получившаяся таблица была преобразована в Excel-файл.

При просмотре получившейся базы данных в Excel обнаружилось, что у большой части авторов не определился пол, поэтому для каждого неопределенного автора был сделан запрос в поисковую систему. Несмотря на персональные запросы, у трёх авторов не удалось определить пол, это иностранные сотрудники. Уже после проделанных последующих заданий и просмотре базы данных у меня возникла идея о возможном преобразовании и улучшении кода для определения пола: выяснилось, что пол автора не определился в том числе у тех авторов, у которых фамилия или все ФИО было написано заглавными буквами. Поэтому можно сделать цикл, который преобразовывает ФИО в нижний регистр и проверяет окончания, сделать можно это с помощью функции «lower()» в Python.

# **Анализ журнала за выбранный период с 2015 по 2017 год**

За исследуемый период было опубликовано 243 статьи. Количество статей резко увеличилось в 2017 году - более чем в 3 раза по сравнению с предыдущим годом (см. Таблицу 2 и Рис. 3). В этот же год журнал перешел в более высокий квартиль – третий, а также увеличилось количество статей, написанных с зарубежными авторами на 4,27%. Несмотря на такие положительные изменения, скачок в количестве опубликованных статей в 2017 году, на мой взгляд, обусловлен другими причинами. Предположу, что не все статьи индексировались в предыдущие года, в связи с чем возник «недоучёт» статей.

Таблица 2. Количество опубликованных статей в журнале за период и по годам.

|  |  |
| --- | --- |
| Период | Количество статей |
| 2015 | 40 |
| 2016 | 45 |
| 2017 | 158 |
| Весь период | 243 |

Рис. 3. Динамика количества статей за 2015-2017 гг.

* Количество авторов.

Следующим этапом был подсчет количества авторов по годам и среднего количества авторов на одну статью (см. Таблицу 3). Видно, что был перекос в большую сторону во все года периода в количестве женщин в общем и на статью, однако заметна тенденция на увеличение доли мужчин в авторстве.

Таблица 3. Общее и среднее количество авторов на статью за период и по годам

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | За период (2015-2017гг.) | За 2015г. | За 2016г. | За 2017г. |
| Общее кол-во авторов | Всего | 402 (399 с определенным полом) | 63 | 76 | 260 |
| Мужчин | 162 | 18 | 24 | 120 |
| Женщин | 237 | 45 | 52 | 140 |
| Среднее кол-во авторов | На одну статью | 1,65 | 1,575 | 1,69 | 1,67 |
| Мужчин на одну статью | 0,54 | 0,29 | 0,32 | 0,46 |
| Женщин на одну статью | 0,8 | 0,71 | 0,68 | 0,54 |

Рис. 4. Динамика среднего количества авторов мужчин и женщин на одну статью за период с 2015 по 2017 год

* Доля женщин среди авторов.

Доля женщин среди авторов уменьшалась с каждыми годом.

Таблица 4 – Доля женщин среди авторов за весь период и по годам

|  |  |
| --- | --- |
| Год | Доля женщин среди авторов |
| 2015 | 0,71 |
| 2016 | 0,68 |
| 2017 | 0,53 |
| Весь период | 0,59 |

Рис. 5. Изменение доли женщин среди авторов с 2015 по 2017 гг.

* Создание дополнительного признака.

Каждой статье была присвоена дополнительная характеристика, которая отражает гендерный состав авторов статьи: все авторы – мужчины, все авторы – женщины или смешанный гендерный состав авторов, переменные m\_only, f\_only и mixed соответственно. Так как всего есть три статьи с неизвестным полом автора из 243 (доля таких статей мала), было принято решение не выделять отдельный тип для этой категории. Поэтому 3 статьи с неопределенным автором входят в категорию mixed, то есть смешанного типа статей. Из таблицы 5 и графика 6 видно, что к 2017 году происходит взаимно компенсирующее сближение долей статей с авторами только мужчинами и только женщинами. На мой взгляд, в 2017 году был улучшен подсчет количества статей и авторов, как уже говорилось выше.

Таблица 5 – Доли типов статей по гендерному составу за весь период и по годам

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Типы статей | За весь период | 2015 | 2016 | 2017 |
| Смешанный | 0,24 | 0,3 | 0,2 | 0,23 |
| Авторы - мужчины | 0,27 | 0,05 | 0,18 | 0,36 |
| Авторы - женщины | 0,49 | 0,65 | 0,6 | 0,41 |

Рис. 6. Динамика изменения долей различных типов статей за период с 2015 по 2017 годы

* Расчет среднего количества цитирований на статью.

Видно, что количество цитирований у типов статей смешанного типа и с авторами только мужчинами была 0, но присутствовало малое цитирование у статей с авторами только женщинами. В 2017 году видна резкая положительная динамика для всех типов. Я связываю это с тем, что в 2017 году возросло кол-во статей, написанных с участием зарубежных авторов, а также в 2017 году журнал перешел в третий квартиль. Однако скорее всего явление резкого увеличения числа цитирований на статью зависит от других факторов, связанных с учетом цитируемости.

Также видно, что цитируемость статей авторов-мужчин выше, чем женских в 2017 году и в целом за период, однако разница в 2017 году несущественная, в то время как за весь период разница в среднем количестве цитирований на статью среди статей, написанных только мужчинами и только женщинами, составляет 0,16.

Таблица 6 – Среднее количество цитирований на статью по типам статьи за весь период и по годам

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Среднее кол-во цитирований на статью | | | |
| Тип статьи | За весь период | 2015 | 2016 | 2017 |
| Смешанный | 0,21 | 0 | 0 | 0,33 |
| Авторы – только мужчины | 0,45 | 0 | 0 | 0,526 |
| Авторы – только женщины | 0,29 | 0,038 | 0,037 | 0,492 |

Рис. 7. Динамика среднего количества цитирований на статью по типам статей с 2015 по 2017 года

* Медианное значение количества цитирований по типам статьи.

Рассчитанные медианные значения оказались равны 0, что означает малую долю цитируемых статей. Связано это скорее всего со особенностью выбранного периода и самим журналом (мало цитируемых статей).

Таблица 7 - Медианное значение количества цитирований по типам статьи за весь период и по годам

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Медианное значение количества цитирований по типам статьи | | | |
| Тип статьи\Период | За весь период | 2015 | 2016 | 2017 |
| Смешанный | 0 | 0 | 0 | 0 |
| Авторы – только мужчины | 0 | 0 | 0 | 0 |
| Авторы – только женщины | 0 | 0 | 0 | 0 |

* 10 самых цитируемых статей.

Среди 10 самых цитируемых статей есть те, что писались женщинами, причем 5 из самых цитируемых статей были написаны исключительно женщинами и одна статья в смешанном составе. Соответственно, 4 статьи написаны только мужчинами. Также среди 5 самых цитируемых статей: 3 из 5 статей написаны исключительно женщинами, 2 - мужчинами. Также примечательно, что все самые цитируемые статьи написаны в 2017 году, а наибольшее количество цитирований равно 4 у статьи, написанной только авторами-мужчинами.

Таблица 8 – 10 наиболее цитируемых статейИзображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

* Альтернативный показатель авторства.

Был посчитан альтернативный показатель авторства методом дробного счета, то есть определен вклад каждого автора в статью. Новая колонка была добавлена в исходную таблицу. Затем посчитана доля женщин среди авторов по данному методу (Таблица 9). Получившиеся показатели доли женщин среди авторов оказались ниже, чем при расчётах с помощью обычного метода. Я связываю данное различие с использованием нового метода, оценивающего именно вклад человека в статью.

Таблица 9 – Доля женщин среди авторов по методы «дробного счета» за период и по годам

|  |  |
| --- | --- |
| Год | Доля женщин среди авторов по методу “дробного счёта” |
| 2015 | 0,71 |
| 2016 | 0,56 |
| 2017 | 0,43 |
| Весь период | 0,5 |

Изображение выглядит как текст, снимок экрана, число, График

Автоматически созданное описание

Рис. 8. Динамика доли женщин среди авторов по методу «дробного счета» за период 2015-2017 гг.

* Среднее количество цитирований по статьям, написанным одним автором и статьям, написанным в соавторстве.

Среднее кол-во цитирований по статьям, написанным одним автором, и статьям, написанным в соавторстве, различается (Таблица 10). Статьи с одним автором цитируются немного чаще, чем статьи в соавторстве, разница в 0,08 цитирований.

Также была посчитана корреляция данных показателей, оказалось, что явных сильных зависимостей нет. Я решила исследовать 2017 год отдельно, так как именно в этом году больше всего цитируемых статей, в то время как в предыдущих двух годах – это значение близко к 0. Кроме того, интересно заметить, что корреляция во всех исследуемых случаях, хоть и не является большой, но имеет отрицательный знак (Таблица 11), а также, что абсолютное значение корреляции для цитируемости и количества женщин в авторстве всегда превышает корреляцию цитируемости и количества мужчин.

Таблица 10 - Среднее количество цитирований по статьям за весь период

|  |  |
| --- | --- |
|  | Среднее количество цитирований по статьям за весь период |
| Один автор | 0,2 |
| Соавторство | 0,12 |

Таблица 11 – Корреляция цитируемости и количеством авторов в целом и по гендерам за весь период и за 2017 год

|  |  |  |
| --- | --- | --- |
| Период |  | Корреляция |
| Весь период | Кол-во авторов и цитируемость | -0,07 |
| Кол-во мужчин и цитируемость | -0,02 |
| Кол-во женщин и цитируемость | -0,08 |
| 2017 год | Кол-во авторов и цитируемость | -0,08 |
| Кол-во мужчин и цитируемость | -0,03 |
| Кол-во женщин и цитируемость | -0,04 |

* Цитируемость без соавторства.

Были выделены статьи, написанные один автором. Их 135. Цитируемость в среднем по статьям, написанным мужчиной, составила 0,19, а женщиной – 0,17. Различие несущественное и может не быть свидетельством гендерного неравенства.

* Публикация в соавторстве.

С помощью цикла в Python были подсчитаны количества мужчин в соавторстве, женщин в соавторстве, мужчин, опубликовавших статью в одиночку, и женщин, опубликовавших статью в одиночку. Затем количество мужчин, опубликовавших статьи в соавторстве, было разделено на общее количество мужчин; такая же процедура была проделана для авторов-женщин. По результатам расчетов получилось, что 46% мужчин публикуются в «СоцИс» в соавторстве, в то время как 32% женщин публикуются в соавторстве в этом же журнале в данном периоде. Следовательно, мужчины публикуются в соавторстве чаще, чем женщины. На мой взгляд, данная тенденция необычна; скорее всего влияет специфика журнала – социология.

* Среднее количество цитирований в расчете на статью по методу «дробного счета».

Также по методу «дробного счета» было найдено среднее количество цитирований на статью по каждому типу статьи в зависимости от гендерного авторского состава. Результаты представлены в таблице 12. Существенные различия между группами видны только в 2017 году: здесь появилось среднее кол-во цитирований на статью у смешанного типа и «авторов – только мужчин», причем в 2017 году у последнего типа больше значение, чем для «авторов – только женщин», что соотносится по тенденциям с обычным методом расчетов. Стоит заметить, что при данном методе расчетов различия в среднем количестве цитирований на статью в группах «авторах-мужчинах» и «авторах-женщинах» чуть более существенно и составляет всего 0,06 вместо 0,04 при обычном методе. Несмотря на это, за весь период видно, что статьи, написанные мужчинами, цитируются больше, чем те, которые написаны женщинами, на 0,16 в среднем. Есть различия с предыдущим методом, описанным ранее, в получившихся значениях в силу использования дробного метода подсчёта среднего количества статей. Расчетные показатели методом «дробного счета» немного меньше, чем у обычного метода.

Таблица 12 - Среднее количество цитирований в расчете на статью за весь период и по годам по методу «дробного счета».

|  |  |  |  |
| --- | --- | --- | --- |
| Год\Тип статьи | Смешанный | Авторы - мужчины | Авторы - женщины |
| 2015 | 0 | 0 | 0,04 |
| 2016 | 0 | 0 | 0,04 |
| 2017 | 0,13 | 0,47 | 0,41 |
| Весь период | 0,08 | 0,37 | 0,21 |

Рис. 9. Динамика изменения среднего количества цитирований в расчете на статью по методу «дробного счета» за 2015-2017 гг.

* Медианное значение количества цитирований по типам статьей.

Медианные значения количества цитирований по типам статей, как и при предыдущем методе, равны нулю. Связано это с малым числом количества цитируемых статей.

Таблица 13 - Медианное значение количества цитирований по типам статьи за весь период и по годам по методу «дробного счета».

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Медианное значение количества цитирований по типам статьи | | | |
| Тип статьи\Период | 2015 | 2016 | 2017 | Весь период |
| Смешанный | 0 | 0 | 0 | 0 |
| Авторы – только мужчины | 0 | 0 | 0 | 0 |
| Авторы – только женщины | 0 | 0 | 0 | 0 |

# **Выводы**

В данной работе были проанализированы такие показатели, как доли мужчин и женщин в авторстве (женщин оказалось в целом больше, чем мужчин), доли типов статей по гендерному составу авторов (на протяжении всего исследуемого периода, несмотря на отрицательную тенденцию, доля статей, написанных исключительно женщинами, была наибольшей по сравнению с двумя другими типами), среднее количество цитирований на статью двумя методами (оказалось, что среднее количество цитирований у статей, написанных авторами-мужчинами больше, чем тех, которые написали женщины, несмотря на большее количество женщин в составе журнала), медианные значения количества статей (которые во всех периодах равны 0 из-за малого количества цитируемых статей). Стоит отметить, что при подсчетах с помощью метода «дробного счета» значения всех показателей получались ниже, чем при обычном. Также было посчитано среднее количество цитирований по статьям, написанным одним автором и соавторстве. В результате было видно, что статьи, написанные одним автором цитируются немного чаще, чем те, что написаны в соавторстве. Было выявлено, что мужчины публикуются чаще в соавторстве. Также стоит учитывать специфику журнала и учитывать особенности выбранного для изучения периода.

Что касается определения гендерного разрыва в академической сфере, то для данного журнала, думаю, требуется дополнительный, более детальный анализ публикаций в последние годы. Из проведенного мной исследования можно выявить следующее противоречие: несмотря на то, что количество женщин в журнале превышает количество мужчин, средняя цитируемость на статью среди статей, написанных исключительно мужчинами, выше, чем у тех, которые написаны женщинами, но разница не является большой, а также стоит помнить, что в данном периоде у «СоцИса» в принципе мало цитирований.