

Эконометрика. Лекция 13

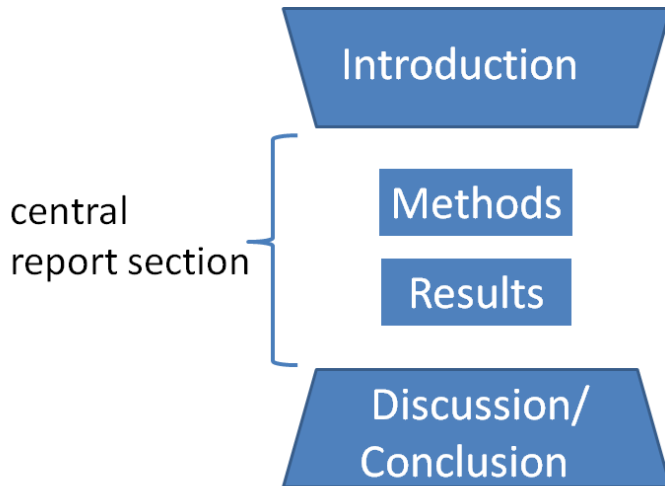
Отчеты о проектах и защита проектов. Структура и оформление

Д. С. Терещенко

НИУ ВШЭ, Санкт-Петербург

16 декабря 2022 г.

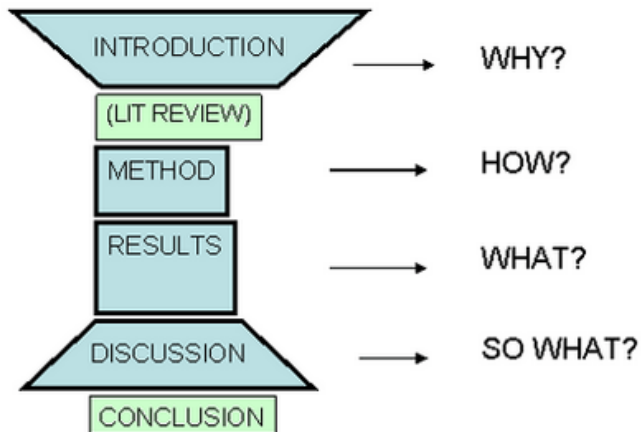
IMRaD — общий подход к структуре научных отчетов



IMRaD — общий подход к структуре научных отчетов

Эконометрика.
Лекция 13

Д. С.
Терещенко



О структуре
отчетов и
презентаций

О защите
проектов

Об оформлении
презентаций

Еще раз о том, что может содержать отчет (и презентация)

- ...
- Вводную часть, описывающую постановку и обоснование вопроса
- Описание данных, разведанализ данных
- Формулировку и обоснование модели или моделей
- Ожидаемые результаты
- Результаты регрессионного анализа
- Анализ результатов, включая результаты проведения статистических тестов или других необходимых процедур
- Ответ на содержательный вопрос в рамках проведенного анализа, при необходимости сопровождаемый уместной визуализацией результатов или аналитическими выкладками на основе проведенного анализа
- Критический анализ полученных результатов, анализ ограничений исследования
- Дискуссионную часть с предложениями по возможному расширению исследования
- Заключение
- ...

Еще раз о том, что может включать описание данных

- **Название базы данных, источник, временной период, включенный в анализ.**
- **Что является единицей наблюдения? Сколько наблюдений имеется?**
- **Преимущества и ограничения используемой базы данных.**
- **Описание все манипуляций, проводимых с данными: слияние двух или нескольких баз; создание новых переменных.**
- **Таблица с описательными статистиками.** При необходимости можно представить описательные статистики по подгруппам.
- **Графики.**

Еще раз о том, что могут включать заключительная часть работы

- Содержательный ответ на вопрос.
- Насколько экономически осмысленны полученные результаты? Как они соотносятся с реальной жизнью и здравым смыслом? Как они соотносятся с результатами предыдущих исследований?
- Кому могут быть полезны полученные результаты? Как их можно применить на практике?
- Обсуждение ограничений исследования и угроз валидности.
- ...

Еще раз о формате защиты

Формат защиты:

- онлайн
- устный доклад с презентацией в течение 7 минут
- вопросы комиссии и ответы в течение 7 минут

Доклад и презентация, ответы на вопросы:

- на защите необходимо присутствовать всей команде (с включенными камерами)
- доклад делает 1-2 человека от команды
- на вопросы всей команде могут отвечать все/любые члены команды, комиссия также может задать вопрос любому члену команды и требовать ответ от него/неё
- присутствовать на защите других команд нет необходимости

Цели презентации

Целью презентации не является:

- удивить/впечатлить слушателей
- рассказать всё, что вы знаете по теме презентации
- рассказать о каждой мелкой детали своей работы

Цели презентации:

- дать аудитории возможность понять ваши идеи/работу
- убедить аудиторию в важности ваших результатов
- получить обратную связь

Общие советы по защите

Во время презентации ...

- будьте полны энтузиазма
- не читайте со слайда или с бумажки
- уложитесь в положенное время

Общие советы по оформлению презентаций

- Каждый слайд и каждый элемент на слайде должен иметь какой-то смысл.
- Используйте списки или схемы вместо сплошного текста.
- Не используйте экзотические шрифты.
- Делайте заголовки на слайдах.
- Используйте разумный размер шрифта.
- Не используйте в таблицах сокращенные названия переменных. Все таблицы должны легко читаться.
- Округляйте до разумного числа знаков после запятой.
- Используйте один язык.
- Используйте единое оформление графиков.

Таблица — очень плохой пример

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew
breaks	1	54	28.15	13.20	26.0	26.41	11.86	10	70	60	1.24
wool*	2	54	1.50	0.50	1.5	1.50	0.74	1	2	1	0.00
tension*	3	54	2.00	0.82	2.0	2.00	1.48	1	3	2	0.00
	kurtosis		se								
breaks		1.33	1.80								
wool*		-2.04	0.07								
tension*		-1.56	0.11								

Таблица — очень плохой пример

```
Descriptive statistics by group
: Female
: Fifth
  vars  n mean  sd median trimmed  mad  min  max range  skew kurtosis  se  IQR Q0.25 Q0.5 Q0.75
1      1 49 0.49 0.08   0.5     0.5 0.08 0.22 0.64  0.42 -0.56    0.56 0.01 0.1  0.44  0.5  0.55
-----
: Male
: Fifth
  vars  n mean  sd median trimmed  mad  min  max range  skew kurtosis  se  IQR Q0.25 Q0.5 Q0.75
1      1 50 0.51 0.09   0.5     0.51 0.08 0.32 0.69  0.37 0.04   -0.4 0.01 0.12  0.45  0.5  0.57
-----
: Female
: Sixth
  vars  n mean  sd median trimmed  mad  min  max range  skew kurtosis  se  IQR Q0.25 Q0.5 Q0.75
1      1 47 0.48 0.09   0.49   0.48 0.1 0.3 0.67  0.37 -0.07   -0.9 0.01 0.13  0.41 0.49  0.54
-----
: Male
: Sixth
  vars  n mean  sd median trimmed  mad  min  max range  skew kurtosis  se  IQR Q0.25 Q0.5 Q0.75
1      1 40  0.5 0.11   0.5     0.5 0.09 0.27 0.76  0.49 0.18   -0.05 0.02 0.13  0.44  0.5  0.57
> |
```

Таблица — плохой пример

	<i>2010</i>	<i>2011</i>	<i>2012</i>	<i>2013</i>	<i>2014</i>	<i>2015</i>	<i>2016</i>	<i>2017</i>
<i>EBIT</i>	6153	6303	6875	8029	8277	9461	9692	9321
<i>T</i>	38,90%	39,30%	38,60%	38,90%	39,50%	39,30%	38,40%	19,80%
<i>EBIT(1-T)</i>	3759	3826	4221	4906	5008	5743	5970	7475
<i>D</i>	1469	1568	1753	1870	1931	2092	2475	2660
<i>ΔNWC</i>	1399	2	-576	3838	-2944	253	-2417	-4211
<i>CAPEX</i>	1464	1276	1478	1930	2125	2332	2224	1918
<i>FCFF</i>	2365	4116	5072	1008	7758	5250	8638	12428

Таблица — хороший пример¹

Наименование переменных	Кол-во наблюдений	Среднее	Стандартное отклонение	Мин.	Макс.
Выручка от реализации, тыс. руб.	808	50 802,53	99 390,48	44	137 8474
Выручка от реализации продукции растениеводства, тыс. руб.	746	39 271,88	64 146,72	8	467 083
Выручка от реализации продукции животноводства, тыс. руб.	369	17 135,72	46 981,82	1	693 309
Материальные затраты, тыс. руб.	808	35 781,95	75 013,96	19	1 099 991
Основные средства (без учета доходных вложений в материальные ценности), тыс. руб.	808	84 996,6	180 361,1	172	2 242 940
Общая земельная площадь, га	808	7843,60	10 260,86	100	100 974

¹Вообще-то это тоже скриншот, т. к. это таблица из чужого исследования. Обладая исходными данными, можно сделать «настоящую» таблицу.

Еще раз об оформлении регрессионных таблиц

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, 91(5), 1369-1401.

TABLE 2—OLS REGRESSIONS

	Whole world (1)	Base sample (2)	Whole world (3)	Whole world (4)	Base sample (5)	Base sample (6)	Whole world (7)	Base sample (8)
	Dependent variable is log GDP per capita in 1995						Dependent variable is log output per worker in 1988	
Average protection against expropriation risk, 1985–1995	0.54 (0.04)	0.52 (0.06)	0.47 (0.06)	0.43 (0.05)	0.47 (0.06)	0.41 (0.06)	0.45 (0.04)	0.46 (0.06)
Latitude			0.89 (0.49)	0.37 (0.51)	1.60 (0.70)	0.92 (0.63)		
Asia dummy				−0.62 (0.19)		−0.60 (0.23)		
Africa dummy				−1.00 (0.15)		−0.90 (0.17)		
“Other” continent dummy				−0.25 (0.20)		−0.04 (0.32)		
R^2	0.62	0.54	0.63	0.73	0.56	0.69	0.55	0.49
Number of observations	110	64	110	110	64	64	108	61

Notes: Dependent variable: columns (1)–(6), log GDP per capita (PPP basis) in 1995, current prices (from the World Bank’s World Development Indicators 1999); columns (7)–(8), log output per worker in 1988 from Hall and Jones (1999). Average protection against expropriation risk is measured on a scale from 0 to 10, where a higher score means more protection against expropriation, averaged over 1985 to 1995, from Political Risk Services. Standard errors are in parentheses. In regressions with continent dummies, the dummy for America is omitted. See Appendix Table A1 for more detailed variable definitions and sources. Of the countries in our base sample, Hall and Jones do not report output per worker in the Bahamas, Ethiopia, and Vietnam.

Еще раз об оформлении регрессионных таблиц

Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, 101(7), 3253–3285.

TABLE 7—SELF-REPORTED VOTE AND NTV IN 1999, SURVEY DATA

	Opposed by NTV in 1999		Supported by NTV in 1999	
	Vote for Unity in 1999 (centrist, progovernment) IV probit	Probit	Vote for OVR in 1999 (centrist, opposition) IV probit	Probit
<i>Panel A</i>				
Watched NTV in 1999	−0.831 [0.301]***	−0.139 [0.128]	1.180 [0.477]**	0.135 [0.156]
Marginal effect	−0.26 [0.09]***	−0.05 [0.04]	0.25 [0.14]*	0.02 [0.02]
Controls	Yes	Yes	Yes	Yes
Observations	901	901	901	901
Number of subregions	42	42	42	42
χ^2 statistics for the exclusion of NTV1999 in the first stage	34.72		24.79	
Supported by NTV in 1999				
	Vote for SPS in 1999 (liberal) IV probit	Probit	Vote for Yabloko in 1999 (liberal) IV probit	Probit
<i>Panel B</i>				
Watched NTV in 1999	1.210 [0.405]***	0.272 [0.149]*	0.467 [0.555]	0.039 [0.178]
Marginal effect	0.24 [0.11]**	0.04 [0.02]*	0.06 [0.08]	0.004 [0.02]
Controls	Yes	Yes	Yes	Yes
Observations	901	901	901	901
Number of subregions	42	42	42	42
χ^2 statistics for the exclusion of NTV1999 in the first stage	28.90		30.47	

Еще раз об оформлении регрессионных таблиц

Lipson, S. K., Zhou, S., Wagner III, B., Beck, K., & Eisenberg, D. (2016). Major differences: Variations in undergraduate and graduate student mental health and treatment utilization across academic disciplines. *Journal of College Student Psychotherapy*, 30(1), 23-41.

Table 4. Multivariable Correlates of Mental Health Problems Among Master's Students, Logistic Regressions.

	Dep			Anx			SI			NSSI			Any			Tx		
	OR	p	95% CI	OR	p	95% CI	OR	p	95% CI	OR	p	95% CI	OR	p	95% CI	OR	p	95% CI
SOC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
HUM	2.15	.001	1.35 3.42	0.98	.94	0.55 1.73	2.10	.01	1.19 3.72	1.58	.03	1.05 2.37	1.70	.003	1.20 2.41	1.02	.95	0.61 1.68
NAT	1.22	.33	0.82 1.82	0.76	.33	0.43 1.33	1.77	.03	1.07 2.94	1.12	.57	0.76 1.66	1.10	.54	0.81 1.49	1.13	.64	0.67 1.91
ART	1.95	<.001	1.39 2.72	1.44	.12	0.91 2.27	1.61	.05	1.01 2.56	1.32	.13	0.93 1.87	1.80	<.001	1.39 2.33	1.09	.71	0.70 1.70
ENG	1.35	.08	0.97 1.89	0.62	.11	0.34 1.11	1.07	.80	0.66 1.71	1.12	.54	0.77 1.63	1.15	.32	0.88 1.49	0.54	.02	0.33 0.89
BUS	0.90	.55	0.64 1.27	0.71	.10	0.48 1.06	0.49	.003	0.31 0.78	0.47	<.001	0.34 0.66	0.68	.002	0.53 0.87	0.66	.07	0.42 1.03
LAW	0.79	.52	0.40 1.59	1.20	.76	0.38 3.79	0.58	.38	0.17 1.96	0.44	.14	0.15 1.31	0.64	.15	0.34 1.18	0.55	.32	0.17 1.81
SW	0.88	.66	0.49 1.56	1.10	.69	0.69 1.73	0.71	.30	0.36 1.37	0.87	.49	0.59 1.29	0.86	.40	0.60 1.23	2.13	.004	1.27 3.56
PH	0.72	.15	0.47 1.13	0.79	.36	0.48 1.31	0.78	.43	0.43 1.43	0.53	.01	0.33 0.83	0.60	.001	0.44 0.82	1.30	.37	0.74 2.30
NUR	0.82	.47	0.47 1.41	1.05	.88	0.55 2.00	0.42	.06	0.17 1.04	0.37	.01	0.19 0.74	0.66	.05	0.44 1.01	0.93	.87	0.40 2.17
MED	1.62	.05	1.00 2.64	1.35	.31	0.75 2.43	0.98	.97	0.44 2.21	0.68	.23	0.36 1.28	1.06	.77	0.71 1.58	1.13	.73	0.57 2.23
OTH	1.07	.68	0.77 1.50	0.99	.95	0.68 1.43	0.90	.62	0.60 1.36	0.70	.02	0.52 0.94	0.95	.67	0.76 1.20	0.94	.75	0.64 1.38
MULT	1.75	.001	1.24 2.47	1.39	.13	0.91 2.14	1.34	.18	0.87 2.07	0.92	.59	0.67 1.25	1.32	.03	1.02 1.70	0.95	.80	0.64 1.41
Constant	0.17	<.001	0.10 0.29	0.09	<.001	0.04 0.19	0.06	<.001	0.02 0.15	0.35	<.001	0.20 0.60	0.54	.004	0.35 0.82	0.22	<.001	0.10 0.47
N	9,282			6,866			9,275			9,171			9,157			2,346		

Note. CI = confidence interval; OR = odds ratio; Dep = Depression; Anxiety (Anx) excludes 2013; SI = suicidal ideation; NSSI = nonsuicidal self-injury; Any = Any Mental Health Problem; Treatment Utilization (Tx) is among students with at least one apparent mental health (MH) problem; SOC = Social Sciences; HUM = Humanities; NAT = Natural Sciences; ART = Art & Design; ENG = Engineering; BUS = Business; LAW = Law; SW = Social Work; PH = Public Health; NUR = Nursing; MED = Medicine; OTH = Other; MULT = Multidisciplinary. All models control for survey year and students' age, gender, citizenship, race/ethnicity, and parental education. Reference category is *Social Sciences*.

Еще раз об оформлении регрессионных таблиц

Kouvonen, A., Kivimäki, M., Virtanen, M., Pentti, J., & Vahtera, J. (2005). Work stress, smoking status, and smoking intensity: an observational study of 46 190 employees. *Journal of Epidemiology & Community Health*, 59(1), 63-69.

Table 2 Relation of smoking status with job strain and effort-reward imbalance: adjusted odds ratios (ORs) and their 95% confidence intervals (CIs)

	Women			Men			p for sex interaction
	Number	OR*	[95% CI]	Number	OR*	[95% CI]	
Job control (component of job strain)							0.776
High job control	11033	1.00		2739	1.00		
Intermediate job control	12428	0.94	(0.87 to 1.01)	2925	0.98	(0.86 to 1.11)	
Low job control	11454	0.96	(0.89 to 1.04)	2490	1.01	(0.88 to 1.17)	
Job demands (component of job strain)							0.616
Low job demands	8547	1.00		2403	1.00		
Intermediate job demands	13825	0.92	(0.85 to 0.99)	3431	0.97	(0.86 to 1.10)	
High job demands	12452	1.01	(0.94 to 1.09)	2310	0.99	(0.87 to 1.14)	
Job strain							0.083
Low strain	8316	1.00		2173	1.00		
Active jobs	10284	1.00	(0.92 to 1.09)	2398	0.87	(0.76 to 1.01)	
Passive jobs	7387	1.02	(0.93 to 1.11)	2021	0.88	(0.75 to 1.02)	
High strain	8792	1.04	(0.96 to 1.13)	1548	1.00	(0.86 to 1.17)	
Effort (component of effort-reward imbalance)							0.019
High effort	12908	1.00		2169	1.00		
Intermediate effort	18037	0.88	(0.83 to 0.94)	4363	1.09	(0.96 to 1.24)	
Low effort	3860	0.83	(0.75 to 0.92)	1605	0.95	(0.81 to 1.11)	
Rewards (component of effort-reward imbalance)							0.292
High rewards	10236	1.00		2405	1.00		
Intermediate rewards	10379	1.03	(0.96 to 1.11)	2355	1.05	(0.91 to 1.21)	
Low rewards	8534	1.20	(1.11 to 1.30)	2501	1.15	(1.00 to 1.33)	
Effort-reward imbalance							0.129
Low imbalance	9044	1.00		2652	1.00		
Intermediate imbalance	11156	1.13	(1.05 to 1.22)	2441	1.01	(0.88 to 1.15)	
High imbalance	8831	1.28	(1.19 to 1.39)	2152	1.13	(0.98 to 1.29)	

*Adjusted for age, basic education, occupational status, type of employment, and marital status.