

How to write an introduction

Anders Munk-Nielsen

October 14, 2024

Why care about writing the introduction? Many economists incorrectly assume that the introduction writes itself once the results are there. This is in spite of the fact that the introduction is the first thing that a reader meets and thus gives the all-important first impression of the text.

In this paper, I will explain some dos and don'ts of writing a good introduction. These are fully subjective and based on nothing more than my rough opinion. I will give examples and name some of the standard paragraphs you can use. My **research question** is: *How does one write a good introduction?*

To answer this, I have read a lot of papers and graded even more student papers. And recently, I started trying to use my system of paragraph archetypes when reading the introductions of papers that I enjoyed reading. This proved to me the strength of the framework, and hopefully it will to you as well by the end.

The challenge in learning to write an introduction is the same as in reading. You have to train yourself to start seeing patterns in what at first seems like endlessly different writings. In the case of introductions, a good place to start is to identify and name the types of paragraphs that people use and to learn what role they play, when to use which ones, and where to put them. A typical introduction consists of something like 5-12 paragraphs, and about 72.3% of those paragraphs come from a finite list of archetypes.

The main contributions of this paper therefore are two-fold: first, to open your eyes when *reading* papers. And second, to improve the quality and structure when you are *writing*, by providing a scaffold. As always, once you have mastered the scaffold and climbed to new heights of insight, you are free to once again kick away the scaffold and laugh at how infantile it is. And as most parents, it will then recognize the rejection as the highest praise.

The closest paper to this is probably Nikolov [2022]. At least from a very quick search. Nikolov provides a long and great paper, which is comprehensive and all. But mine is short, concise, and written in a funny style.

Another strand of literature has looked at this other aspect of writing. Here, my contribution is something else.

The rest of the paper is organized as follows: Section 1 gives all the details. Enjoy.

1 General writing tips

- **Use short and simple words, always:** Monosyllabic lexemes compound per-spicuity and assuage opacity.
- **Use active language:** “I assume”, “I use data on”, etc. Avoid “It is assumed”.
 - Use “I” for single-authored work and “we” with multiple co-authors.

1.1 How to write paragraphs of text

A paragraph of text is concerned with a single “message.” You can say several facts in a paragraph, but they should all be aligned towards a single goal. For instance, “*I am able to overcome the challenges faced in prior work for three reasons. First, ...*” – this paragraph will contain three facts, but the first sentence shows what the three are pointing towards.

The first sentence of the paragraph prepares the reader for what is coming. If I read the first sentences of each paragraph in the introduction, I want to understand where all the information is (the research question, the methods/data used, the key takeaways, the most closely related paper, ...).

- *The findings mentioned above may seem surprising, but they align with what has been found in a related literature on school choice...*
 - *The data I use is ideal to study my question for three reasons.*
 - *In studying this question, the key empirical challenge one must overcome is to handle the endogenous selection into ownership.*

These first sentences do not explain anything, but they explain the logic of why the paragraph comes now, and prepare the reader for what type of information is coming. This allows the reader to be conduct *active reading* – for example, a reader may be patiently reading all the while waiting for you to address one particular concern which is especially important to this reader.

Before and after writing your introduction, create a **bullet list** with one bullet per paragraph. In 3-5 words, what is that paragraph? If you find you have to write a lot, consider if your paragraph is doing too much? Then use the list to make sure that you are not repeating yourself.

2 Paragraph Archetypes

A list of paragraph archetypes and the purpose they serve

- **Why care:** The first paragraph – the 'hook' that gets the reader interested. This is the most important paragraph. Experiment with many different approaches and pick the best one.
- **In this paper:** Clearly states what is done, and prepares the reader for what to expect in terms of methodology.
- **To do this:** Explains methodology.
- **The key challenge / contribution:** What is the central thing that you overcome which has held researchers from addressing your interesting research question for so long? It can be a novel dataset, new identification strategy, a reform you can exploit, a new model, etc.
- **My data is ideal because:** This paragraph is sometimes *The key challenge is*, but when it is not, it may serve an independent purpose because it allows you to flesh out the virtues of your dataset, *if it has relevance* for your key contribution (e.g. you observe a variable that is normally missing; you have panel as opposed to just cross sectional data; etc.).
 - Don't insert a full-length data description into your introduction unless it is relevant (i.e. plays an important role for your relative contribution over the state-of-the-art).

- **The closest paper** (also called *I contribute to a large literature on...*): Shows what the state-of-the-art is, and should also explain what your contribution is relative to that. (Also points out an obvious referee.)
 - **Do:** *Prior research has found that [...]. This paper shares [...] but is the first to [...].*
- **Another strand of literature** that you contribute to: secondary contribution.
- **I find that:** Highlight the most central result. Remember, *tell me only what I absolutely need to know!* There must be a narrative hook for a result to make it to the introduction. (A valid reason is that it's your research question.)

3 Why care? The first paragraph

The very first paragraph should answer “*why care?*” It should provide a **hook** for your reader to catch their interest. Inspired by Nikolov [2022], a good hook is either

1. A puzzle,
2. A tradeoff,
3. Some data that shows a development / trend / change that is surprising / worrying / etc.
4. A controversial statement: “*Economists have belatedly come to recognize that the ‘law of one price’ is no law at all.*” (Varian, 1980)
5. A compelling story.

The best hooks simultaneously answer the *why care*-question and illustrate a key mechanism or tradeoff at play. Often, this can be done by clearly stating the research question of the paper, which should contain both.

The hardest part to write is the first sentence. Popular choices include:

1. The research question: great to just state this up front,
2. *Economists have long realized that...*: Emphasizes how your topic is smack in the middle of economics,

Some examples

- **Varian (1980):** *“Economists have belatedly come to recognize that the ‘law of one price’ is no law at all. Most retail markets are instead characterized by a rather large degree of price dispersion. [...]”*

⇒ Starts with a controversial (somewhat funny) statement and a fact that flies in the face of the state-of-the-art.

- **Rust (1986):** *“IN THE DURABLE GOODS LITERATURE, two issues have generated controversy over the last decade: (i) does the existence of a competitive secondary market limit profits of a monopolist seller in the primary market? (ii) does the monopolist’s choice of built-in durability depart from the competitive or socially optimal level? The “conventional wisdom” on these issues is that [...]”*

⇒ two research questions.

- **Shapiro (2018):** *“How does television advertising affect the consumer choice problem? After a consumer watches a commercial, internalizes its message, and decides a product is desirable, she must take further action to obtain the product. With groceries, she must go to the supermarket. With many consumer products, a computer with internet will allow the consumer to make the purchase. With prescription drugs, the consumer must go to the physician to obtain a prescription and then to the pharmacy to purchase the drug. With many steps between the advertising incidence and purchase, at some stage of the process, the consumer might well choose a product different from the one advertised. This decision may be due to difficulty in remembering advertisements, agency problems in obtaining products, or simply that advertising convinces a consumer to go to a retailer, computer, or physician. In short, an advertisement could affect the choice process without leading the consumer to buy the advertised product.*

In this paper, I identify the existence of positive spillovers of television advertising in the market for antidepressants. Given these spillovers, I construct and estimate a demand model that allows such spillovers. To quantify the potential size of the incentive effects of spillovers on firm behavior, I conduct a supply-side analysis supposing that the firms are able to decide advertising jointly, and I compare this outcome to a benchmark competitive outcome in which firms optimally free ride in the antidepressant market.”

1. The research question (how “advertising → purchase”)
 2. How “advertising → purchase” plays out/is interfered with in *different* markets, ending with *his* market.
⇒ *foreshadows the key empirical challenges*
 3. Next paragraph: *In this paper...* where he lays out the agenda for the entirety of the paper.
- **Einav, Finkelstein & Cullen (2010; QJE):** *“The welfare loss from selection in private insurance markets is a classic result in economic theory. It provides, among other things, the textbook economic rationale for the near-ubiquitous government intervention in insurance markets. Yet there has been relatively little empirical work devoted to quantifying the inefficiency that selection causes in a particular insurance market, or the welfare consequences of potential policy interventions in that market. This presumably reflects not a lack of interest in this important topic, but rather the considerable challenges posed by empirical welfare analysis in markets with hidden information.”*
 1. *Economists have long thought about the welfare loss from selection*
 2. Puzzle = little empirical work.
 3. Controversial (?) statement = methodological challenges have held research back
(implicitly: will be overcome)
 - **Brancaccio, Kalouptsi and Papageorgiou (2020; ECTA):** *“More than 80% of international trade is carried by the global shipping industry. To export, an exporter has to find an available vessel and contract for a voyage and a price. In turn, the ship is optimally choosing its travels in search of cargo, thinking about its future options. This spatial equilibrium between exporters and ships determines the trade costs (i.e. transportation prices) that different countries face, as well as the trade flows between different regions of the world. What is the role of geography (i.e. country locations, natural inheritance of goods) in determining trade costs and flows? Is the matching process between exporters and ships efficient? How does ship behavior affect the behavior of exporters and the resulting trade flows? In this paper, we use detailed micro-data on vessel movements [...”*

Establishes:

1. The industry is important
 2. The key ingredients in the structural model (exporters find ships, ships choose travel destinations and exporters, spatial equilibrium determines trade costs → trade flows)
 3. The research questions
- **Armstrong (2016; JEEA):** *“Consider a consumer who wishes to purchase one product from the several variants that are available. In some cases she might know exactly what she wants in advance, and no prior market investigation is needed. In other situations, she might be so ill-informed that she stumbles randomly from one option to another, incurring a search cost each time, until she discovers something good enough. Between these extremes, though, a consumer [...]*

In this paper I say that a product is “prominent” for a consumer when it is inspected early in that consumer’s chosen order of search. The theory outlined in the next section shows that a product is more likely to be prominent for a rational consumer if that product: (i) has a distribution for match utility that is better (in the sense of first-order stochastic dominance) or riskier; (ii) is expected, or observed, to have a lower price, or (iii) is less costly to inspect. These three factors are discussed in more detail in the following.”

1. A compelling story
2. In this paper...

4 In this paper

Very quickly, you have to explain what happens in this paper. This should give an early hint to the reader about what to expect and what broad category your paper fits into. It should clarify

- Whether the paper is empirical (and if so, broadly what data is used)
- What the overall methodology is (structural/reduced form/etc.)
- [optional] A teaser for the main results / answer to research question

If you haven’t stated the research question in the first paragraph, it should come either *In this paper* or immediately after. You don’t have to actually state it as a question,

but it doesn't hurt. And if I had a dollar for every time I had thought to myself "*This time I stated the research question in blazing letters, clear to everyone*", only to realize that it wasn't there...

5 The key challenge / contribution

Very early, you should explain to your reader what the main obstacle in your research has been. Typically, there's a particular source of endogeneity that everyone's worried about in your part of the discipline. Ideally, your main contribution is that you overcome the main challenge in your literature.

For example, maybe the key challenge is that of selection, and your main contribution is that you study a reform that circumvents this by controlling uptake. Then you have to make your reader appreciate how hard this problem is, and give some intuition for why you have a cure for the disease. And either then or in a paragraph in the near future, you should provide references to influential papers that have been struggling with this issue before you (or have succeeded in spite of ignoring this very dangerous issue), and how you are doing a much better job of dealing with it because of reasons. But you can also circle back to this when you get to your *The closest papers*-paragraph.

6 Related literature

This part can be either a separate section or a few paragraphs in the introduction. Regardless, the structure is: a sequence of paragraphs that each deal with a strand of the literature, and where each paragraph ends with *My/Our contribution to this strand of literature is [...]*. You can also state your contribution in the middle, so long as you use the exact words [my/our] "*contribution is to [...]*" for each paragraph.

For fairly broad strands of literature, it can be beneficial to write "*Within this literature, the closest two papers to my paper are X and Y.*" and then highlight what they did and what your relative contribution is.

The purpose of the literature review is mainly to contextualize your contribution, although it also makes the editor's job of finding suitable referees easier. So make sure that each paragraph serves one of these purposes. Try to avoid "*many people have looked at this question before*"-paragraphs. Instead think about *how* this question was investigated before, and in what way your approach/data/model contributes. Ideally,

you think of 2-4 key contributions and organize your 'strands of the literature' according to the contributions. For instance, your contributions might be: *i)* providing a causal estimate of an effect ("prior work used IV, RD, or DiD, but you have an experiment"), *ii)* estimating the effect in a given context ("I also contribute to a wider literature on pharmaceutical demand"), and *iii)* analyzing a specific policy ("... environmental car tax policies ...").

7 Examples

7.1 Ursu 2018

The abstract of Ursu (2018) does it all.

- *"Online search intermediaries, such as Amazon or Expedia, use rankings (ordered lists) to present third-party sellers' products to consumers. These rankings decrease consumer search costs and increase the probability of a match with a seller, ultimately increasing consumer welfare. [research question] Constructing relevant rankings requires understanding their causal effect on consumer choices. [challenge] However, this is challenging because rankings are endogenous: consumers pay more attention to highly ranked products, and intermediaries rank the most relevant products at the top. **In this paper, I** use the first dataset with [overcome this] experimental variation in the ranking from a field experiment at Expedia to make three **contributions**. **First**, I identify the causal effect of rankings and show that they affect what consumers search, but conditional on search, do not affect purchases. **Second**, I quantify the effect of rankings using a sequential search model and find an average position effect of \$1.92, which is lower than literature estimates obtained without experimental variation. I also use model predictions, data patterns, and a feature of the data set (opaque offers) to show rankings lower search costs, instead of affecting consumer expectations or utility. **Finally**, I show a utility-based ranking built on this model's estimates benefits consumers and the search intermediary."*

References

Plamen Nikolov. Writing tips for economics research papers – 2021-2022 edition. *IZA DP No. 15057*, 2022.