

# Writing

## Seminario de Tesis PEG

### Econ 4600

Ignacio Sarmiento-Barbieri

Universidad de los Andes

January 26, 2026

# Agenda

① Review: Forma y contenido

② Secciones de un artículo de investigación publicable

- Portada
- Introducción
- Literatura previa
- Middle Bits
- Conclusiones
- Bibliografía

③ Sobre escribir

# Agenda

① Review: Forma y contenido

② Secciones de un artículo de investigación publicable

- Portada
- Introducción
- Literatura previa
- Middle Bits
- Conclusiones
- Bibliografía

③ Sobre escribir

# Estructura

- ▶ El trabajo del economista es argumentar (McCloskey, 1985)
- ▶ Forma y contenido **NO** son separables.
  - 1 La investigación es un ejercicio en persuasión. Sus lectores son la audiencia. Trate de interesarlos en su tesis y convencerlos de su argumento.
  - 2 El contenido de su trabajo es evaluado en primera instancia por la manera en que lo presenta. Si el trabajo está mal escrito el lector lo va a leer con escepticismo.
  - 3 “Ningún buen paper, sin importar cuán bien construido, brillante y bien escrito, surgió por primera vez de la impresora del autor en esa forma. Reescribir es el verdadero arte de escribir.” (Goldin and Katz, 2009)

# Estructura

- ▶ Asuma un lector (muy) impaciente → Entender la distinción de leer por gusto y por necesidad.
- ▶ Hay que organizar el paper en estilo “triangular” o de “periódico”, no en estilo de “novela”
  - ▶ Los periódicos comienzan con la parte más importante, luego completan los antecedentes para los lectores que continuaron y querían más detalles.
  - ▶ Un buen chiste o una novela de misterio tiene un largo desarrollo hasta el remate (*punchline*) al final.
  - ▶ Usted está escribiendo un paper no una novela : ponga el *punchline* justo al frente y luego explique lentamente el chiste. Los lectores no se quedan para encontrar el mensaje en la Tabla 12.

# Como lo transmitimos?

- ▶ Documentos escritos
- ▶ Presentaciones

# Agenda

① Review: Forma y contenido

② Secciones de un artículo de investigación publicable

- Portada
- Introducción
- Literatura previa
- Middle Bits
- Conclusiones
- Bibliografía

③ Sobre escribir

# Portada

- 1 Título**
- 2 Autor(es)**
- 3 Abstract/Resumen**
- 4 Clasificación JEL (Journal of Economic Literature) (2-5 códigos)**
- 5 Palabras clave (2-5 palabras)**
- 6 Pie de página**
  - ▶ Información de contacto de los autores (email, universidad...).
  - ▶ Conflicto de interés y financiación.
  - ▶ Agradecimientos.

# Portada

## Housing Discrimination and the Toxics Exposure Gap in the United States: Evidence from the Rental Market

Peter Christensen, Ignacio Sarmiento-Barbieri and Christopher Timmins\*

August 16, 2020

### Abstract

Local pollution exposures disproportionately impact minority households, but the root causes remain unclear. This study conducts a correspondence experiment on a major online housing platform to test whether housing discrimination constrains minority access to housing options in markets with significant sources of airborne chemical toxics. We find that renters with African American or Hispanic/LatinX names are 41% less likely than renters with White names to receive responses for properties in low-exposure locations. We find no evidence of discriminatory constraints in high-exposure locations, indicating that discrimination increases relative access to housing choices at elevated exposure risk.

**Key words:** Housing Discrimination, Correspondence Experiment, Air Toxics

**JEL Classification:** Q51, Q53, R310

\*Christensen: Department of Agricultural and Consumer Economics, University of Illinois, 431 Mumford Hall, 1301 W. Gregory Urbana Illinois 61801 (email: pchrist@illinois.edu); Sarmiento-Barbieri: Department of Economics, University of Los Andes, Calle 19A No. 1-37 Este Bloque W, Bogota, Colombia (email: lsarmiento@uniandes.edu.co); Timmins: 209 Social Sciences Building, P.O. Box 90097, Durham, North Carolina 27708 (email: christopher.timmins@duke.edu). We thank David Albouy, Spencer Banzhaf, Pat Bayer, Jonathan Colmer, Rebecca Diamond, Nick Kuminoff, Juan Carlos Serrato, Daniel Sullivan, Reed Walker, Randy Walsh, and seminar participants at NBER Energy and Environmental Economics, The Workshop on Environmental Economics and Data Science, Triangle Resources and Environmental Economics (TREE) Seminar, UC Berkeley, UC Santa Cruz, UC San Diego, UC Santa Barbara Occasional Workshop in Environmental Economics, UNLV, University of Miami, University of Virginia, and Wharton School for excellent comments. We thank Davis Berlind, Christopher Kim, Daniel Rychel, Tom Phillips and student assistants in the University of Illinois Big Data and Environmental Economics and Policy (BDEEP) Group and the Duke Environmental Justice Lab for research assistance. We acknowledge generous support from the Russell Sage Foundation, National Science Foundation, and the National Center for Supercomputing Applications. The experiment was registered on the AEA RCT Registry as trial 3366 (Christensen et al., 2018) and the human subjects protocol for this research design was approved by the University of Illinois Institutional Review Board (IRB #18381) on 12/07/2017. All data and code will be available prior to publication using DOI [10.5281/zenodo.3492234](https://doi.org/10.5281/zenodo.3492234). All errors are our own.

# Primer Filtro: Título

- ▶ Dado que el lector es (muy) impaciente es importante capturar la atención rápido
- ▶ El título → fundamental.
- ▶ Hay varias estrategias
  - ▶ "Cute" titles
    - ▶ "Star wars: The empirics strike back"
    - ▶ "Banks as Potentially Crooked Secret-Keepers"
    - ▶ "Are Residential Electricity Prices Too High or Too Low? Or Both?"
  - ▶ Detailed titles

# The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges.\*

Marcela Eslava<sup>†</sup>      John Haltiwanger<sup>‡</sup>      Nicolas Urdaneta <sup>§</sup>

February 10, 2023

# Primer Filtro: Título

- ▶ Dado que el lector es (muy) impaciente es importante capturar la atención rápido
- ▶ El título → fundamental.
- ▶ Hay varias estrategias
  - ▶ "Cute" titles
    - ▶ "Star wars: The empirics strike back"
    - ▶ "Banks as Potentially Crooked Secret-Keepers"
    - ▶ "Are Residential Electricity Prices Too High or Too Low? Or Both?"
  - ▶ Detailed titles
  - ▶ Something in between
  - ▶ Or one word

THE  
**QUARTERLY JOURNAL  
OF ECONOMICS**

---

Vol. 138

2023

Issue 3

---

AI-TOCRACY\*

MARTIN BERAJA  
ANDREW KAO  
DAVID Y. YANG  
NOAM YUCHTMAN

# Primer Filtro: Título

## Elephants

*By MICHAEL KREMER AND CHARLES MORCOM\**

*Many open-access resources, such as elephants, are used to produce storable goods. Anticipated future scarcity of these resources will increase current prices and poaching. This implies that, for given initial conditions, there may be rational expectations equilibria leading to both extinction and survival. The cheapest way for governments to eliminate extinction equilibria may be to commit to tough antipoaching measures if the population falls below a threshold. For governments without credibility, the cheapest way to eliminate extinction equilibria may be to accumulate a sufficient stockpile of the storable good and threaten to sell it should the population fall. (JEL Q20)*

# Primer Filtro: Título

## DAMS\*

ESTHER DUFLO AND ROHINI PANDE

This paper studies the productivity and distributional effects of large irrigation dams in India. Our instrumental variable estimates exploit the fact that river gradient affects a district's suitability for dams. In districts located downstream from a dam, agricultural production increases, and vulnerability to rainfall shocks declines. In contrast, agricultural production shows an insignificant increase in the district where the dam is located but its volatility increases. Rural poverty declines in downstream districts but increases in the district where the dam is built, suggesting that neither markets nor state institutions have alleviated the adverse distributional impacts of dam construction.

## Segundo Filtro: Abstract

- ▶ La mayoría de las revistas permiten entre 100 y 300 palabras. Obedezca este límite desde el inicio.
- ▶ La función principal del resumen es comunicar la contribución central y novedosa de su artículo
- ▶ Describa el contenido sin formulas o tecnicismos innecesarios.
- ▶ Di lo que encuentras, no lo que buscas. No escribas “se analizan los datos, se prueban los teoremas, se discute...”

# Segundo Filtro: Abstract

Abstract: Example. Luco (2019)

## Who Benefits from Information Disclosure? The Case of Retail Gasoline<sup>†</sup>

By FERNANDO LUCO\*

*How does online price disclosure affect competition when both consumers and firms can use the disclosed information? This paper addresses this question exploiting the sequential implementation of an online price-disclosure policy in the Chilean retail gasoline industry. The results show that disclosure increased margins by 9 percent on average, though the effects varied across the country depending on the intensity of local search behavior. Because margins increased the least, and even decreased, in high-search areas, where income is also higher, the results also show that price disclosure policies may have important distributional effects. (JEL D83, L11, L71, L81, O13, Q35, Q41)*

# Tercer Filtro (quizás el más importante): Introducción

- ▶ El un lector (muy) impaciente, si el título y el abstract capturaron la atención van a seguir por la intro
  - ▶ Tercer filtro de lectores.
  - ▶ Primeros tres/cuarto párrafos son claves.
  - ▶ Ideal comunicar su argumento desde el inicio (pregunta, respuesta, posicionamiento).
  - ▶ Existe una formula?

# The Introduction Formula by Keith Head

- ▶ Hook
- ▶ Question
- ▶ Antecedents
- ▶ Value-Added
- ▶ Road map

# The Introduction Formula

## Hook

- ▶ Attract the reader's interest by telling them that this paper relates to something interesting.
- ▶ What makes a topic interesting? Some combination of the following attributes makes Y something worth looking at.
  - ▶ Y matters: When Y rises or falls, people are hurt or helped.
  - ▶ Y is puzzling: it defies easy explanation.
  - ▶ Y is controversial: some argue one thing while other say another.
  - ▶ Y is big (like the service sector) or common (like traffic jams).
  - ▶ A story hook (perhaps relating to a common activity or an issue that affects a lot of people)

# The Introduction Formula

## Hook

- ▶ Things to avoid:
  - ▶ The bait and switch : promising an interesting topic but delivering something else, in particular, something boring.
  - ▶ “all my friends are doing it” : presenting no other motivation for a topic than that other people have written papers on it.
  - ▶ The reader will be much more motivated to read the rest of the paper if you challenge his or her intuition right from the get-go.
  - ▶ Your readers are your audience. They have better things to do than read your paper. Make them interested in your thesis and convinced of your argument in the first two paragraphs.

# The Introduction Formula

Hook: Examples I. Oster (2012)

## 1. Introduction

Five to ten percent of adults in Sub-Saharan Africa are infected with the human immunodeficiency virus (HIV) and the primary mode of transmission in the region is heterosexual sex. For this reason, sexual behavior change is a major focus of HIV prevention efforts and understanding changes in behavior is important both for predicting the future path of the epidemic and for developing policy.

# The Introduction Formula

Hook: Examples II (Albouy, 2020)

## 1. Introduction

Economic theory leans heavily on the idea that goods may be complements in consumption. While the joint demand for private goods has been studied extensively, little has been said about the joint demand for public goods. Studying the joint demand for public goods is difficult as individuals cannot purchase them directly, but only indirectly, such as through housing. To the best of our knowledge, no study has estimated the joint demand for public

# The Introduction Formula

Hook: Examples III (Burlon et al., 2024)

In recent years, the demand for digital means of payment for retail purposes has steadily increased, while the use of cash for transactions has gradually declined (Auer, Cornelli, and Frost 2020). In response to this shift in payment technologies and preferences, central banks all over the world have started to investigate the potential benefits and implications of issuing central bank digital currencies (CBDCs). The ultimate goal of introducing a CBDC is to ensure that individuals operating in an increasingly digitalized economy keep having access to the safest form of money: central bank money. The most discussed challenge of issuing a CBDC is the risk of bank disintermediation as households substitute bank deposits for CBDC, thereby reducing a relatively cheap funding source for banks.

# The Introduction Formula

Hook: Examples IV (Brüggemann, 2021)

The taxation of top income earners is a controversial topic. In public debates over recent years, supporters of raising marginal tax rates on top income earners usually have the intention to close fiscal deficits and/or decrease economic inequality. Opponents of this view instead demand lower rates on top incomes as a means of shifting the tax burden away from high-income, high-productivity households and boosting economic activity. An increasing number of academic papers has studied the optimal level of top marginal tax rates (TMTRs). Spurred by Diamond and Saez's (2011) recommendation to impose high marginal tax rates on top income earners of up to 80 percent, a recent wave of quantitative studies, which I discuss in further detail below, uses dynamic general equilibrium models to determine optimal TMTRs. The results differ widely depending on specific modeling choices, especially regarding households' labor income processes and the implied labor supply elasticities among top income earners.

None of these papers features entrepreneurs. This paper closes that gap in the literature by explicitly modeling entrepreneurship based on Cagetti and De Nardi (2006) to evaluate the level and economic impact of optimal top marginal tax rates. Accounting for entrepreneurs is important for several reasons. First, in

# The Introduction Formula

## Question

- ▶ Tell the reader what this paper actually does.
- ▶ Think of this as the point in a trial where having detailed the crime, you now identify a perpetrator and promise to provide a persuasive case.
- ▶ The reader should have an idea of a clean research question that will have a more or less satisfactory answer by the end of the paper.
- ▶ The question may take two paragraphs. At the end of the first (2nd paragraph of the paper) or possibly beginning of the second (3rd paragraph overall) you should have the “This paper addresses the question” sentence.

# The Introduction Formula

Question (Eslava et al., 2024)

## 1 Introduction

A prevalent feature of market economies is heterogeneity of firm and establishment size, growth, and a host of establishment attributes correlated with size (e.g., productivity, exports, survival). What are the sources of such heterogeneity, how does the answer matter for aggregate productivity and welfare? A crucial insight from the macro misallocation literature is that there are wedges (often referred to as distortions) impacting establishment size relative to what would be implied by establishment true productivity, and that this leads to aggregate productivity losses, especially in developing economies. Contributions in trade and IO have focused on how firm/establishment size is impacted by attributes such as demand (quality/appeal), markups, or costs, finding that idiosyncratic demand-side factors dominate.<sup>1</sup>

How do these findings relate to each other? Do wedges lie mainly on the cost or demand sides? What sources of heterogeneity across productive units are most harmful for aggregate activity and which are most enhancing, and how is that harm reflected in the size distribution of firms? We examine these questions by developing a unified conceptual, measurement, and estimation structure that accounts for a uniquely rich set of establishment (plant) attributes, and taking it to detailed data on manufacturing establishments. Our framework takes advantage of data on output and input prices and quantities to measure and estimate the role of these detailed establishment attributes. We consider establishment-level quality shifters, markups, and two distinct dimensions of idiosyncratic marginal costs: technical efficiency and quality-adjusted input prices, including wages, material prices, and, in an extension, idiosyncratic user cost of capital inclusive of factor-biased distortions. Residual wedges help account for the differences between the size distribution implied by theory and the data even after incorporating all of the components separately measured in our analysis.<sup>2</sup>

# The Introduction Formula

## Antecedents-Literature

- ▶ More on this latter

# The Introduction Formula

## Road-map

- ▶ Outline the organization of the paper.
- ▶ Avoid writing an outline so generic that it could apply to any paper (“the next section is the middle of the paper and then we have the end”).
- ▶ Instead customize the road map to the project and possibly mention pivotal “landmarks” (problems, solutions, results...) that will be seen along the way.
- ▶ But keep this short because many readers will now be eager to get to the heart of the paper.

*This paper proceeds as follows. The following section provides background on the experimental design and sample. Section 3 discusses results on the discrimination- exposure relationship by toxic concentration and by distance to TRI facility. Section 4 discusses heterogeneity in the discrimination-exposure relationship by price and housing/neighborhood characteristics. Section 5 concludes.*

# The Introduction Formula

Road-map (Currie et al., 2015)

The rest of the paper proceeds as follows: Section I presents an analytical framework which helps motivate the empirical analysis. Section II discusses the data, and Section III discusses the research design. Sections IV and V outline the econometric specifications and results for housing values and infant health respectively. Finally, Section VI interprets the results, and Section VII concludes.

# The Introduction Formula

## General Advice

- ▶ Write the intro first but then read and edit it every time you compose other parts of the paper.
- ▶ Thus by the end, the intro will have received more attention, more times, than any other part of the paper.
- ▶ The introduction is not just important because of the “first impressions” idea that it will tilt the referee for or against you (though it probably will).
- ▶ It is also vital to making sure you know yourself what you are doing in the paper and why.
- ▶ If you can’t write a good introduction, then you may be writing the wrong paper.

# Como discutir la literatura previa

(o si debo tener una sección aparte)

- ▶ Dos escuelas:

- 1 Parte de la introducción (posicionamiento y contribución).

- 2 Sección independiente.

- ▶ Integre su discusión de la literatura previa bajo el hilo común de trabajos anteriores en relación con su tesis principal.

# Como discutir la literatura previa

- ▶ Al escribir esta sección tengan en cuenta su objetivo y que esta sección tiene dos funciones
  - ▶ La primera es simplemente para demostrar su familiaridad con el tema. Hay que proporcionar una síntesis de lo que ha leído, rastrear el desarrollo de temas importantes y extraer cualquier tensión en la investigación previa.
  - ▶ La segunda función es sentar las bases para su artículo y preparar al lector con respecto a lo que su documento trae a la mesa en términos de novedades.
- ▶ En algunos casos, estas dos funciones están en direcciones opuestas: la primera hacia incluir tantas fuentes como sea posible, la segunda hacia seleccionar solo aquellas que son útil para su argumento.
- ▶ Recuerde: su principal contribución será juzgada en relación con trabajos de economía anteriores y la disciplina económica!

# Como discutir la literatura previa

## Antecedents

- ▶ Identify the prior work that is critical for understanding the contribution this paper will make.
- ▶ The key mistake to avoid here are discussing papers that are not essential parts of the intellectual narrative leading up to your own paper.
- ▶ Give credit where due but establish, in a non-insulting way, that the prior work is incomplete or otherwise deficient in some important way.

# Como discutir la literatura previa

Parte de la introducción: Value-Added

- ▶ Describe approximately 3 contributions this paper will make relative to the antecedents.
- ▶ This paragraph might be the most important one for convincing your judges
- ▶ A big difference between it and the earlier “question” paragraph is that the contributions should make sense only in light of prior work whereas the basic research question of the paper should be understandable simply in terms of knowing the topic (from the hook paragraph).
- ▶ “Antecedents” and “Value-added” may be intertwined.

# Como discutir la literatura previa

Parte de la introducción (posicionamiento y contribución).

## ***Full-Length Research Articles***

These articles are typically empirical in nature and deal with an important issue within the field of education economics. They should be divided into numbered sections (1., 2., 3...) and subsections (1.1.1, 1.1.2,...), beginning with an introduction that clearly states the purpose of the study while avoiding a detailed description of the previous literature. In addition to the introduction, full-length research articles typically include a background section, a methods section and/or a data section, a results section, and a section that concludes. Appendices should be identified as using capital letters (Appendix A, Appendix B...), and equations, tables and figures that appear in the appendices should be given separate numbering (e.g., Table A.1, Table A.2,...).

# Como discutir la literatura previa

## Ejemplos

- ▶ Parte de la introducción (posicionamiento y contribución).
  - ▶ Luco (2019)
  - ▶ Eslava et al. (2024)
  - ▶ Baker et al. (2016)
- ▶ Sección independiente.
  - ▶ Brüggemann (2021)
  - ▶ Burlon et al. (2024)

# Middle Bits

- ▶ Depends on the type of paper one writes
- ▶ Papers usually have these sections
  - ▶ Theoretical Framework (applied micro, macro, econometrics)
  - ▶ Institutional Context (applied micro, macro)
  - ▶ Data and Descriptive Statistics (applied micro, macro)
  - ▶ Empirical Framework (applied micro, macro)
  - ▶ Simulation, Calibration (macro, econometric )
  - ▶ Results (applied micro, macro)
  - ▶ Application (macro, econometric )

# Marco Teórico

- ▶ Primitivas: ¿Cómo son las preferencias y/o la tecnología?
- ▶ Variables y parámetros
- ▶ Supuestos
- ▶ Equilibrio
- ▶ Problema de maximización y condiciones de primer orden
- ▶ Predicción comprobable
- ▶ Pruebas

# Marco Teórico

## Example I (Albouy et al., 2020)

### 2. A simple model of public good complements

In principle, complementarity in public goods, e.g., warm weather and a community pool, is no less important than between private goods, swimming trunks and goggles. An important difference is that local public goods are bought indirectly through housing. This purchase is developed in the model below.

Preferences are represented by a Cobb-Douglas function: the utility of person  $i$  in location  $j$  is  $U_{ij} = Q_{ij}y^\alpha x^{1-\alpha}$ , where  $y$  is the quantity of the housing good consumed, with price  $v_j$ ,  $x$  is a numeraire good, and  $\alpha \in (0, 1)$  is a fixed parameter.  $Q_{ij}$  gives the value of location  $j$  to person  $i$ , which is log-linear in interacted amenities:

$$\ln Q_{ij} = (\theta^P + \theta^{PH}H_j) P_j + \theta^H H_j + \ln \xi_j + \epsilon_{ij} \quad (1)$$

where  $P_j$  denotes the environmental amenity,  $H_j$  denotes the crime level, and  $\xi_j$  other commonly-valued amenities. The parameter  $\epsilon_{ij}$  is an idiosyncratic taste shock for the neighborhood.

The parameters  $\theta^P > 0$  and  $-\theta^H > 0$  define the base elasticities of willingness-to-pay for the environmental amenity and safety (minus crime), respectively. The interaction parameter  $\theta^{PH}$  describes the complementarity, which we predict to be negative. Alternatively, safety and parks are complements. These terms may be arranged as  $(\theta^H + \theta^{PH}P_j) H_j + \theta^P P_j$  to illustrate how the cost of crime rises when the environmental amenity is higher.<sup>8</sup>

Denote our measure of crime,  $\tilde{H}_j = H_j + a_j$ , where  $a_j$  captures measurement error. This error adds to the unobserved amenity term:  $\tilde{\xi}_j = \xi_j + (\theta^H + \theta^{PH}P_j) a_j$ . Taking these shifts into account, the indirect utility function is given by:

$$\ln U_{ij} = -\alpha \ln v_j + (\theta_j^P + \theta^{PH}H_j) P_j + \theta^H \tilde{H}_j + \tilde{\xi}_j + \epsilon_{ij}.$$

Solving for log housing price, and letting  $V_{ij} = \ln v_{ij}$ , it is natural to separate out the park-crime interaction.

$$\begin{aligned} V_j &= \frac{\theta_j^P}{\alpha} P_j + \frac{\theta^H}{\alpha} \tilde{H}_j + \frac{\theta^{PH}}{\alpha} (P_j \times H_j) + \frac{\tilde{\xi}_j + \epsilon_{ij} - \ln U_{ij}}{\alpha} \\ &\equiv \beta^P P_j + \beta^H \tilde{H}_j + \beta^{PH} (P_j \times \tilde{H}_j) + \xi_j^* + u_{ij} \end{aligned} \quad (2)$$

where  $\beta^k = \theta^k/\alpha, k \in \{P, H, PH\}$ ,  $\xi_j^* = \tilde{\xi}_j/\alpha$ , and  $u_{ij} = (\epsilon_{ij} - \ln U_{ij})/\alpha$ . This specification predicts that  $\beta^P > 0$  and  $\beta^H < 0$ . If parks and safety are complementary, then  $\beta^{PH} < 0$ . This linear model predicts that above a certain level of crime, a park becomes a public bad. If

$$\tilde{H}_j \geq -\frac{\beta^P}{\beta^{PH}} = \frac{\theta^P}{\theta^{PH}}, \quad (3)$$

then households will pay to live away from the park.<sup>9</sup> As shown in Banzhaf (2015) for the case of individual amenities, hedonic estimates that exploit exogenous changes in the level of one or both public goods complements may shift an entire hedonic price function and identify a lower bound on the Hicksian equivalent surplus.

# Marco Teórico

Examples Micro-Theory, Macro, IO

- ▶ Elsinger et al. (2019)
- ▶ Burlon et al. (2024)
- ▶ Eslava et al. (2024)

# Contexto institucional

- ▶ ¿Qué factores institucionales debe conocer el lector para seguir el argumento?
- ▶ Se necesita un balance. Escriba para un lector que:
  - 1 No conoce del tema.
  - 2 Sólo quiere saber lo relevante para entender el punto.
- ▶ Example: see [Luco \(2019\)](#)

# Descripción de los datos

- ▶ Datos:
  - ▶ ¿Cuándo se recolectaron? ¿Dónde? ¿Por qué? ¿Por quién? ¿Cómo se seleccionó la muestra? ¿A quién se entrevistó o cómo se recopilaron los datos? ¿Cuál es el tamaño de la muestra? ¿Cómo se compara con la población de interés? ¿Pierdes alguna observación? ¿Por qué? ¿Tuviste que imputar algún valor y, de ser así, cómo lo hiciste? ¿Existen variables proxies para lo real? ¿Qué mide exactamente cada variable o cómo se construyó?
- ▶ Estadísticas descriptivas: Arme una buena tabla de estadísticas descriptivas. En la narración describa los contenidos de su tabla de estadísticas descriptivas contando una historia sobre ellos, no se limite a escribir una aburrida enumeración de medias.
- ▶ Balance de las muestras: en los casos que la una variable de interés es dicotómica (o categórica),
  - ▶ ¿cómo difieren los grupos de tratamiento y control a lo largo de la media de las variables discutidas en la subsección anterior?

# Descripción de los datos

Ejemplo: Who Benefits from Information Disclosure? The Case of Retail Gasoline (Luco, 2019)

TABLE 1—SUMMARY STATISTICS: MARGINS

	Unit	Mean	Standard deviation
<i>Panel A. Inflation-adjusted margins</i>			
Before disclosure	(CLP \$/liter)	69.26	23.12
After disclosure	(CLP \$/liter)	71.64	22.24
Number of observations	5,795		
Percent of stations			
<i>Panel B. Station characteristics</i>			
Convenience store		42.97	
Pharmacy		4.69	
Public restrooms		37.50	
Repair shop		33.59	
Self-service pumps		17.97	
Open 24 hours		91.41	
Number of requests			
	Per capita (mean)	Level (mean)	Level (standard deviation)
<i>Panel C. Search requests near a station (within a month)</i>			
1 kilometer	0.0005	86.87	199.05
3 kilometers	0.0046	627.95	1,324.03
5 kilometers	0.0113	1,413.59	2,996.35

Note: CLP\$ stands for Chilean pesos.

# Descripción de los datos

Ejemplo: Who Benefits from Information Disclosure? The Case of Retail Gasoline (Luco, 2019)

Table 1 reports summary statistics of the data at the station level. The table is divided into three panels, which differ in the variables of interest. The first panel reports statistics on margins and shows that unconditional mean margins before the implementation of disclosure were equal to 69.25 Chilean pesos per liter. This average increased to 71.64 Chilean pesos per liter once disclosure was implemented. However, as it will be shown below, these averages hide significant heterogeneity across locations.

The second panel summarizes station characteristics and shows that 43 percent of stations had a convenience store, almost 5 percent had a pharmacy, 38 percent had public restrooms, 34 percent had a repair shop, 18 percent offered self-service pumps (in addition to full service), and 91 percent operated continuously. I later use this information to study whether changes in margins may be associated with

increasing differentiation through the website, as stations could use the website to inform consumers about the services they offer.<sup>12</sup>

The third panel summarizes the number of monthly search requests executed in the neighborhood of each gas station, for different distance thresholds. The table shows that, though on average there is little search, there is significant heterogeneity across locations.

# Marco Empírico

## ► Estrategia de identificación:

- ▶ ¿Cómo sería el experimento ideal para estudiar su pregunta? ¿Qué tan cerca estás de ese ideal y qué te impide acercarte?
- ▶ Discuta a su vez cómo su estrategia de identificación (IV, Dif-in-Dif, RDD, Efectos Fijos, PSM...) trata o no con
  - 1 la heterogeneidad no observada,
  - 2 la causalidad reversa o simultaneidad, y
  - 3 el error de medición.
- ▶ Debilidades y fortalezas de su estrategia de identificación ¿Cuáles son los supuestos necesarios para la credibilidad de los resultados?

# Marco Empírico

Ejemplo: The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality (Ananat, 2011)

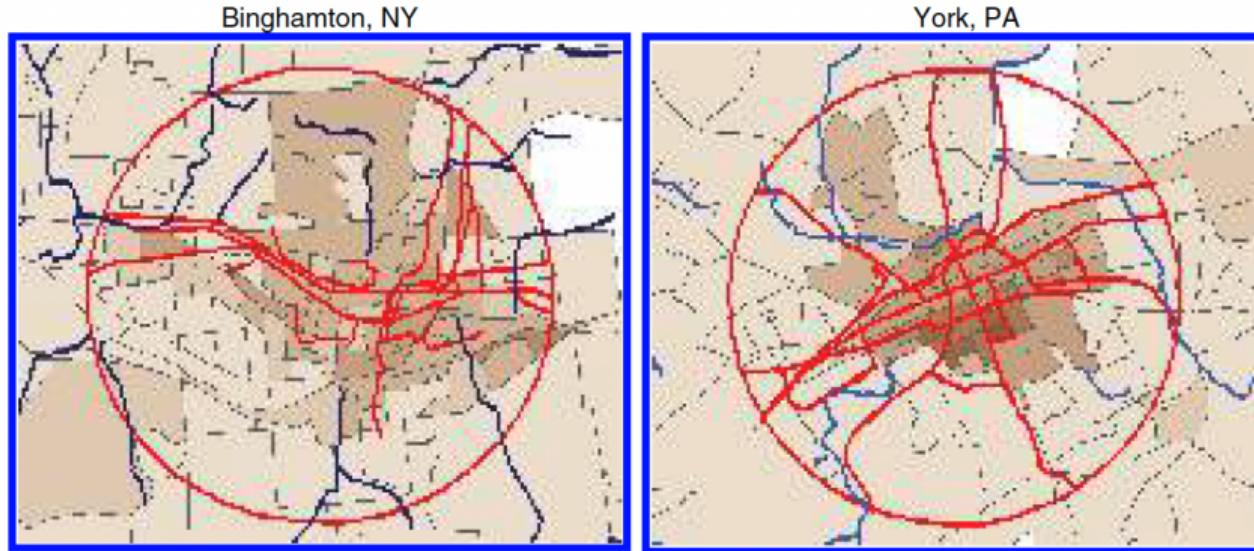


FIGURE 1. THE NATURAL EXPERIMENT—2 EXAMPLES

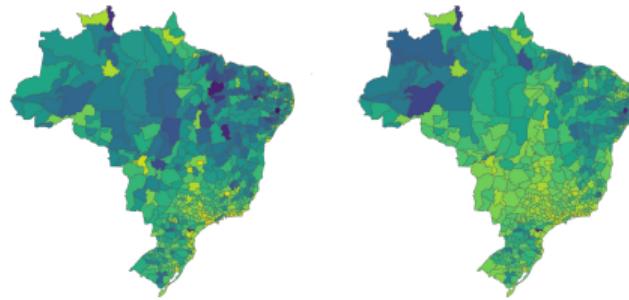
# Marco Empírico

Ejemplo: The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality (Ananat, 2011)

Formally, I approximate the ideal randomized experiment for places by exploiting the configuration of tracks into shapes that define uniform subunits of land (details follow later in this section) in a city's historical center, conditional on total track length. The process of extracting railroad data from a city map is explained in detail in Appendix B. Briefly, I collected from nineteenth-century maps of 121 cities information about the railroads that covered a 50-square kilometer circular area centered on the historical city (the circles are visible in Figure 1).<sup>15</sup> The choice of this area size provides the advantage that, while all the cities studied exceed this perimeter today, about 75 percent of the cities were smaller than that area when mapped. So, for most cities, this measure includes railroads that were laid on unoccupied land without any need to consider human occupants.

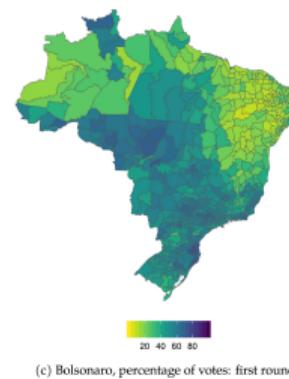
# Marco Empírico

Ejemplo: Economic shocks, gender, and populism: Evidence from Brazil (Barros and Silva, 2025)



(a) Male shock  
0.00 0.25 0.50 0.75

(b) Female shock  
0.00 0.25 0.50 0.75



(c) Bolsonaro, percentage of votes: first round  
20 40 60 80

# Marco Empírico

- ▶ Estrategia de estimación: escriba la ecuación que va a estimar

$$y_{it} = \alpha_i + \beta COVID33_{it} + \theta X_{it} + \delta_t + u_{it} \quad (1)$$

- ▶ Describa todas las variables y clarifique la notación ¿A qué corresponde la variable  $COVID33_{it}$ ?
- ▶ Describa el (los) parámetro(s) de interés y su interpretación.
- ▶ ¿Cómo los estimará?
- ▶ ¿Cómo tratará los errores estándar?

# Marco Empírico

Aside: Diff in Diff

$$y_{it} = \beta D_{it} + \gamma_i + \delta_t + \varepsilon_{it},$$

- ▶  $Y_{it}$ : outcome for unit  $i$  at time  $t$
- ▶  $D_{it}$ : treatment indicator ( $= 1$  if treated *and* post-treatment)
- ▶  $\gamma_i$ : unit fixed effects
- ▶  $\delta_t$ : time fixed effects
- ▶  $\beta$ : DiD estimator (average treatment effect on treated, under assumptions)

# Marco Empírico

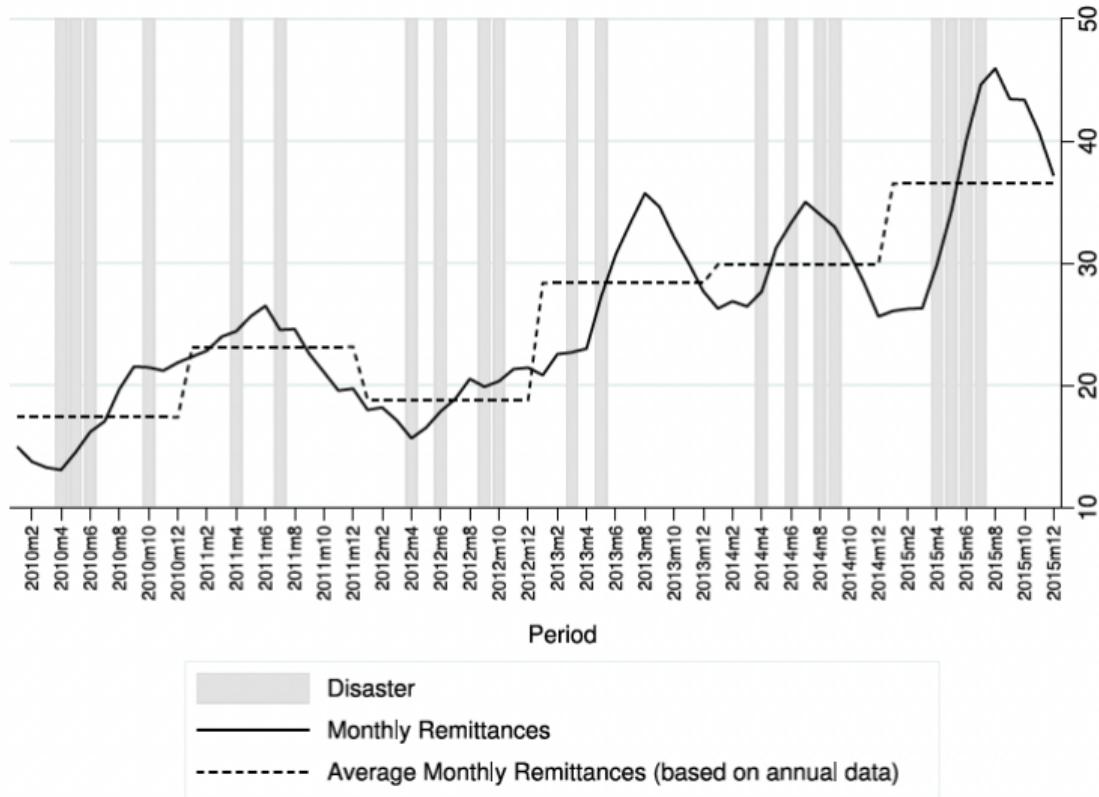
## Aside: Event-Study Specification

$$y_{it} = \underbrace{\sum_{j=-m}^n \gamma_j \cdot D_{i,t-j}}_{\text{Event Study Terms}} + \alpha_i + \delta_t + \beta \cdot X_{it} + \varepsilon_{it}.$$

- ▶  $y_{it}$  is the outcome of interest for unit  $i$  at time  $t$ .
- ▶  $D_{i,t-j}$  is an indicator variable for **event time**  $j$ , meaning the event occurred  $j$  periods before the current observation's calendar time.
- ▶ A separate coefficient  $\gamma_j$  is estimated for each event time  $j$ .
  - ▶ Post-event coefficients ( $\gamma_j$  for  $j \geq 0$ ) capture the dynamic effects of treatment over time.
  - ▶ Pre-event coefficients ( $\gamma_j$  for  $j < 0$ ) serve as a placebo or falsification test. In the absence of anticipation effects or omitted confounders, these pre-event terms should show no trend in  $j$ .
- ▶  $\alpha_i$  and  $\delta_t$  are (panel) fixed effects for units and time, respectively.
- ▶  $\beta \cdot X_{it}$  represents optional control variables.
- ▶ Plotting the estimated  $\gamma_j$  coefficients against  $j$  typically yields the well-known event-study diagram.
- ▶ See [Miller \(2023\)](#)

# Marco Empírico

Ejemplo: Responding to natural disasters: What do monthly remittance data tell us? (Bettin et al., 2025)



# Marco Empírico

Ejemplo: Responding to natural disasters: What do monthly remittance data tell us? (Bettin et al., 2025)

Fig. 2 provides an intuitive illustration of how relevant it could be to consider the monthly dynamics of remittances to identify migrants' response to disasters at home. Here we plot the monthly variability of remittances to Bangladesh overlaid with bars for the timing of disasters in the country and with a broken line for the average yearly amount that Bangladeshi migrants in Italy transfer home, zooming in on the period 2010–2015.<sup>11</sup>

# Marco Empírico

Ejemplo: Responding to natural disasters: What do monthly remittance data tell us? (Bettin et al., 2025)

## 4. Empirical strategy

To estimate the dynamic response of monthly remittance flows to disasters in the home country, we exploit the exogenous nature of such catastrophic events and conduct an event study analysis. We use a non-parametric event study specification similar to Dobkin et al. (2018). One of the main advantages of this approach is that it allows to flexibly observe and describe the pattern of remittance flows relative to the precise time when a disaster occurs.

# Marco Empírico

Ejemplo: Air quality and early-life mortality: Evidence from Indonesia's wildfires (Jayachandran, 2009)

## A. Empirical Model and Outcome Variable

The goal of the empirical analysis is to measure the effect that air pollution from the wildfires had on early-life mortality. Ideally, there would be data on all pregnancies indicating which ended in fetal, infant, or child death, and the following equation would be estimated:

$$(1) \quad Survive_{jt} = \beta_1 Smoke_{jt} + \delta_t + \alpha_j + \varepsilon_{jt}.$$

The variable  $Survive_{jt}$  is the probability that fetuses whose due date is month  $t$  and whose mothers reside at the time of the fires in subdistrict  $j$  survive to a certain point, such as live birth, one year, etc. The prediction is that  $\beta_1$  is negative, or that exposure to smoke around the time of birth reduces the probability of survival.

In practice, mortality records are unavailable for Indonesia, and survey samples are too small to examine the effects of month-to-month fluctuations in pollution. For example, the 2002 Demographic and Health Survey has on average one birth and 0.05 recorded child deaths per district-month for the affected cohorts.<sup>6</sup> Therefore, the approach I take is to infer early-life mortality by measuring “missing children.” The outcome measure is the cohort size for a subdistrict-month calculated from the complete 2000 Census of Population for Indonesia. The estimating equation is

$$\ln(CohortSize)_{jt} = \beta_1 Smoke_{jt} + \beta_2 PrenatalSmoke_{jt} + \beta_3 PostnatalSmoke_{jt} + \delta_t + \alpha_j + \varepsilon_{jt} \quad (2)$$

The dependent variable,  $\ln(CohortSize)_{jt}$ , is the natural logarithm of the number of people born in month  $t$  who are alive and residing in subdistrict  $j$  at the time of the 2000

# Resultados

- ▶ Esta sección es obviamente la sección más importante del documento.
- ▶ Paradójicamente, es quizás también la sección menos leída .
- ▶ La secciones de resultados tienen su propia estructura

# Orden de los resultados

- ▶ Los resultados se presentan desde los más parsimoniosos hasta los menos parsimoniosos
  - ▶ Resultados (No paramétricos?) Preliminares
  - ▶ Resultados (Paramétricos?) principales
  - ▶ Pruebas de robustez
  - ▶ Extensiones, Mecanismos, etc.
  - ▶ Limitaciones

# Resultados (No paramétricos?) Preliminares

Ejemplo: Estimates of the impact of crime risk on property values from Megan's laws (Linden and Rockoff, 2008)

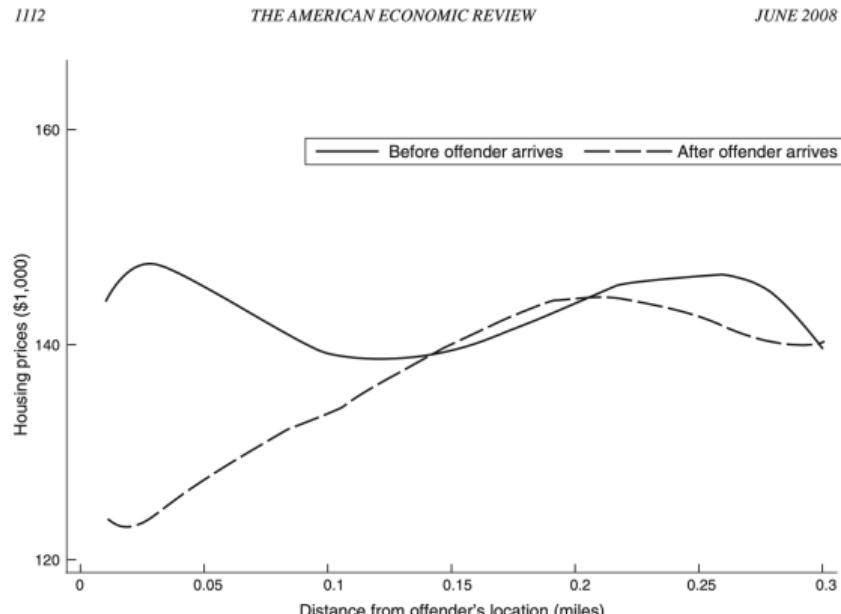


FIGURE 2B. PRICE GRADIENT OF DISTANCE FROM OFFENDER  
(Sales during year before and after arrival)

Note: Results from local polynomial regressions (bandwidth = 0.075 miles) of sale price on distance from offender's future/current location.

# Resultados Principales

1 Comience con su resultado principal.

- ▶ Discuta la “magnitud” del parámetro estimado de interés (significancia económica).
- ▶ Discuta la “precisión” del estimador (significancia estadística).
- ▶ Relacione la “magnitud” del parámetro estimado con lo que encuentra la literatura.
- ▶ ¿El resultado es el esperado? ¿posibles interpretaciones alternativas?

2 Sólo después de “agotar” el análisis de su resultado principal continúe con los resultados secundarios (efectos heterogéneos, etc.).

3 No use adjetivos para describir sus resultados (increíble, espectacular...).

# Resultados principales y tablas de regresión

Ejemplo: Estimates of the impact of crime risk on property values from Megan's laws (Linden and Rockoff, 2008)

- ▶ Las tablas de regresión deben incluir: título descriptivo, numeración (diferentes especificaciones), nombre de la variable dependiente, nombres claros de las variables independientes, otra información relevante.
- ▶ Orden desde los más parsimoniosos hasta los menos parsimoniosos

1116

THE AMERICAN ECONOMIC REVIEW

JUNE 2008

TABLE 3—IMPACT OF SEX OFFENDERS' LOCATIONS ON PROPERTY VALUE AND SALE PROBABILITY

	Log (sale price) pre-arrival		Log (sale price), pre- and post-arrival			Probability of sale†	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Within 0.1 miles of offender	-0.340 (0.052)*	-0.007 (0.013)	-0.007 (0.012)	<0.001 (0.013)	-0.006 (0.012)	-0.006 (0.012)	-0.029 (0.035)
Within 0.1 miles × post-arrival			-0.033 (0.019)+	-0.041 (0.020)*	-0.036 (0.021)+	-0.116 (0.059)+	0.126 (0.059)*
Dist*≤ 0.1 miles × post-arrival (0.1 Miles = 1)						0.107 (0.064)+	
Within 1/3 miles of offender					-0.010 (0.007)		
Within 1/3 miles × post-arrival					0.010 (0.010)	0.003 (0.016)	0.004 (0.016)
H <sub>0</sub> : within 0.1 miles × post-arrival = 0			p-value = 0.079	p-value = 0.0443	p-value = 0.0828	p-value = 0.0502	p-value = 0.0361
Housing characteristics	✓	✓	✓	✓	✓	✓	✓
Year fixed effects	✓						
Neighborhood-year fixed effects	✓	✓	✓	✓			
Offender area-year fixed effects					✓	✓	✓
Restricted to offender areas					✓	✓	✓
2 years pre- and post-arrival							
Standard errors clustered by...	Neighborhood	Neighborhood	Neighborhood	Neighborhood	Offender area	Offender area	Offender area
Sample size	164,993	164,968	169,557	169,557	9,086	9,086	1,519,364
R <sup>2</sup>	0.01	0.84	0.83	0.83	0.75	0.75	0.01

Note: Pre-arrival (post-arrival) refers to the two-year period before (after) the date upon which offenders registered their current address. Standard errors in parentheses.

\* Significant at 5 percent level.

+ Significant at 10 percent level.

† Probability sale is measured as percentage points, e.g., probability of sale + 1 would be 100 percentage points.

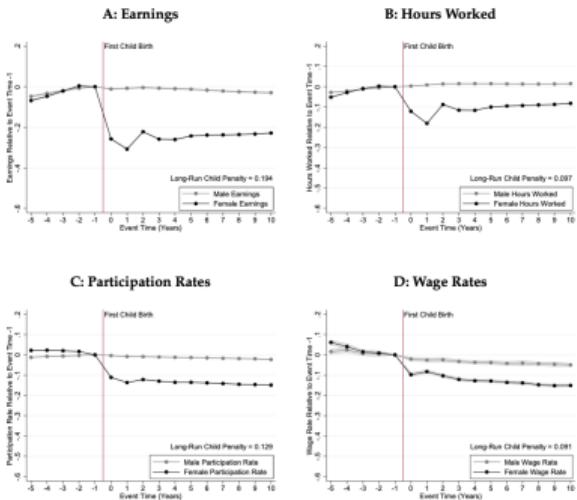


# Resultados (tablas y figuras)

Ejemplo: Children and gender inequality: Evidence from Denmark Kleven et al. (2019)

- Todas las tablas y figuras deben ser auto-contenidas → se debe poder entender su contenido sin ir al texto del artículo.

Figure 3: Impacts of Children



Notes: The graphs show event time coefficients estimated from equation (1) as a percentage of the counterfactual outcome absent children (i.e.,  $P_t^0 = \hat{\alpha}_t^0 / E[\hat{Y}_{t+1}^0 | t]$  as defined in section 3.1) for men and women separately and for different outcomes. Each panel also reports a "child penalty"—the percentage by which women are falling behind men due to children—defined as  $P_t = (\hat{\alpha}_t^{m\bar{x}} - \hat{\alpha}_t^{w\bar{x}}) / E[\hat{Y}_{t+1}^0 | t]$ . The long-run child penalty is measured at event time 10. All of these statistics are estimated on a balanced sample of parents who have their first child between 1985–2002 and who are observed in the data during the entire period between 5 years before and 10 years after child birth. The effects on earnings and participation are estimated unconditional on employment status, while the effects on hours worked and wage rates are estimated conditional on participation. The shaded 95 % confidence intervals are based on robust standard errors.

# Pruebas de robustez

1 ¿Qué tan “robustos” son los resultados a especificaciones alternativas?

- ▶ Selección de muestra.
- ▶ Simulaciones con parámetros diferentes.

2 Tests de placebo.

3 Estructura de rezagos y reversión a la media (panel).

# Extensiones y Limitaciones

- ▶ Extensiones: aquí es donde podría explorar la heterogeneidad del tratamiento, o dividir mi muestra entre hombres y mujeres, rural y urbano, o por industria.
- ▶ Limitaciones: Ningún resultado empírico es perfecto. ¿Cómo se limita la validez interna? ¿Cómo se limita la validez externa? ¿Qué no dicen sus resultados, es decir, qué errores podría cometer la gente al interpretarlos?

# Conclusiones

- ▶ Resumen
  - ▶ Incluya la respuesta a la pregunta de investigación
  - ▶ Incluya lo que usted quiere que el lector recuerde.
  - ▶ No incluya nuevos resultados.
- ▶ Discuta las limitaciones del estudio y como se podrían resolver.
- ▶ Discuta implicaciones de política (si las hay).
- ▶ Implicaciones para investigación futura.

# Bibliografía

- ▶ Incluya únicamente los trabajos que cita en el texto.
- ▶ Orden alfabético.
- ▶ Diferentes estilos posibles para referencias bibliográficas y citas (APA o Chicago son comunes en economía). Lo importante es ser consistente.

# Agenda

① Review: Forma y contenido

② Secciones de un artículo de investigación publicable

- Portada
- Introducción
- Literatura previa
- Middle Bits
- Conclusiones
- Bibliografía

③ Sobre escribir

# Escribir es pensar (McCloskey, 1985)

- 1 Escribir es pensar: uno no aprende los detalles de un argumento hasta que los escribe. En ese proceso se descubren las fallas de fondo en el pensamiento.
- 2 No espere a terminar la investigación para empezar a escribir. La investigación es escribir.
- 3 A nadie le gusta la ansiedad de llenar un página vacía. Supérelo.

# ¿Cómo escribir? (forma)

- 1 "The rule of clearness is not to write so that the reader can understand, but so that he cannot possibly misunderstand" Christopher Morley.
- 2 El texto está compuesto de secciones. Las secciones están compuestas de párrafos. Los párrafos están compuestos de oraciones. Las oraciones están compuestas de palabras.
  - ▶ ¿La sección ayuda a comunicar mejor el argumento? ¿El párrafo ayuda a comunicar mejor el argumento? ¿La oración ayuda a comunicar mejor el argumento?
  - ▶ La mayoría de párrafos tienen muchas oraciones. La mayoría de oraciones tienen muchas palabras. Recorte, recorte y recorte nuevamente.
  - ▶ La repetición aburre. La repetición aburre. La repetición aburre.

# ¿Cómo escribir? (forma)

3 A cada párrafo le corresponde un único punto.

- ▶ Trate de organizar el párrafo en niveles descendientes: de lo general a lo particular.

4 Ayude al lector con una estructura (e.g. signposting.)

- ▶ Los resultados encontrados se resumen en tres puntos: primero...
- ▶ Hay cuatro características del entorno institucional que son particularmente relevantes...

# ¿Cómo escribir? (forma) (McCloskey, 2000, Cochrane, 2005)

- ▶ Keep it short!
  - ▶ Every word must count.
  - ▶ Shorter is better
  - ▶ Don't repeat!
- ▶ Use active tense!
  - ▶ If it is solo authored paper "I" is fine!
- ▶ Use normal sentence structure
  - ▶ subject, verb, object

## ¿Cómo escribir? (forma) (McCloskey, 2000, Cochrane, 2005)

- ▶ Fluency Can Be Achieved by Grit
- ▶ Write Early Rather Than Late
- ▶ Use Your Ear. Read, Out Loud
- ▶ Watch Punctuation
- ▶ Avoid This, That, These, Those

# Bibliografía I

- Albouy, D., Christensen, P., and Sarmiento-Barbieri, I. (2020). Unlocking amenities: Estimating public good complementarity. *Journal of Public Economics*, 182:104110.
- Ananat, E. O. (2011). The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, 3(2):34–66.
- Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.
- Barros, L. and Silva, M. S. (2025). Economic shocks, gender, and populism: Evidence from brazil. *Journal of Development Economics*, 174:103412.
- Bettin, G., Jallow, A., and Zazzaro, A. (2025). Responding to natural disasters: What do monthly remittance data tell us? *Journal of Development Economics*, 174:103413.
- Brüggemann, B. (2021). Higher taxes at the top: The role of entrepreneurs. *American Economic Journal: Macroeconomics*, 13(3):1–36.
- Burlon, L., Muñoz, M. A., and Smets, F. (2024). The optimal quantity of cbdc in a bank-based economy. *American Economic Journal: Macroeconomics*, 16(4):172–217.
- Cochrane, J. H. (2005). Writing tips for ph.d. students.

# Bibliografía II

- Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.
- Elsinger, H., Schmidt-Dengler, P., and Zulehner, C. (2019). Competition in treasury auctions. *American Economic Journal: Microeconomics*, 11(1):157–184.
- Eslava, M., Haltiwanger, J., and Urdaneta, N. (2024). The size and life-cycle growth of plants: The role of productivity, demand, and wedges. *Review of Economic Studies*, 91(1):259–300.
- Goldin, C. and Katz, L. (2009). The ten most important rules for writing your job market paper.
- Jayachandran, S. (2009). Air quality and early-life mortality: Evidence from Indonesia's wildfires. *Journal of Human Resources*, 44(4):916–954.
- Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.
- Linden, L. and Rockoff, J. E. (2008). Estimates of the impact of crime risk on property values from Megan's laws. *American Economic Review*, 98(3):1103–27.
- Luco, F. (2019). Who benefits from information disclosure? the case of retail gasoline. *American Economic Journal: Microeconomics*, 11(2):277–305.
- McCloskey, D. (1985). Economical writing. *Economic Inquiry*, 23(2):187–222.

# Bibliografía III

McCloskey, D. N. (2000). *Economical Writing*. Waveland Pres, Inc., Long Grove, IL.

Miller, D. L. (2023). An introductory guide to event study models. *Journal of Economic Perspectives*, 37(2):203–230.

# Agenda

## ④ Cuestiones Administrativas

# Cuestiones Administrativas

- ▶ Materiales: <https://ignaciomsarmiento.github.io/teaching/Tesis.html>
- ▶ Entregas: Bloque Neón

# Cronograma

Fecha	Actividad
Enero 21 - 28	Clases magistrales: introducción, estructura y escritura de un trabajo de investigación en Economía
Febrero 2 - Febrero 18	Discusión individual con el profesor (obligatoria)
Febrero 20	Entrega: introducción preliminar y estructura de la tesis
Febrero 23 - Marzo 11	Discusión individual con el profesor (optativa)
Marzo 13 6 p.m.	Entrega primer documento
Marzo 16 - 22	Receso
Marzo 27 6 p.m.	Entrega: Referee report
Marzo 30 - Abril 3	Semana Santa
Abril 6 - Abril 8	Clases magistrales: presentación de un trabajo de investigación en Economía
Abril 13 - Abril 29	Presentación (30 min.)
Mayo 4 - Mayo 13	Discusión individual con el profesor (optativa)
Mayo 15 6 p.m.	Entrega documento final

# Espacio de acompañamiento

- ▶ Reunión individual (una obligatoria) con el profesor.
- ▶ La reunión obligatoria se agenda en el sign-up sheet publicado en Bloque Neón.
- ▶ Reuniones adicionales son optativas y se pueden coordinar en el horario de atención o de clase.
- ▶ Se discutirá el estado actual del trabajo y los pasos a seguir.

# Espacio de acompañamiento

- ▶ Reunión obligatoria
- ▶ Slide deck (las reuniones son de 30 min.)
  - ▶ Pregunta de investigación + posicionamiento
  - ▶ Datos
  - ▶ Estrategia de Datos y/o Modelo