Environmental/Energy Economics (ARE 261)
Lecture 1: Intro and Sufficient Statistics

Joseph S. Shapiro

UC Berkeley and NBER

Introduction: What is Environmental Economics, What Will We Cover

- Demand for Environmental Goods (Non-Market Valuation)
  - **Valuation methods: product demand (/hedonics), health, travel demand**, contingent valuation
- Supply of Environmental Goods
  - **Cost/production functions, effects of regulation, market power**, entry
- Broader Public Finance Theory
  - **Externalities, public goods, Tiebout sorting**, social choice theory,
- Other environmental issues
  - **Trade**; Choice of instrument; development; political economy; innovation; energy
- Closely related
  - Resources (exhaustible & renewable); Energy
- Environmental topics
  - **Climate change; pollution (water, air, land)**; resources (water quantity, timber, minerals); protected areas

Motivation: Why Study Environmental Economics

- Socially important
  - Affects 100% of GDP
  - Fix market failures
  - Important time in policy
  - Exciting time in research (Nordhaus Nobel this week!)

- Excellent setting to learn about economics
  - Public: externalities, public goods, benefit-cost analysis
  - Regulated industries: excellent data
  - Fixing market failures: no equity-efficiency tradeoff
  - Young field of economics

- Good area as Ph.D. student
  - Young field of economics
  - Regulated industries: excellent data
  - Fixing market failures: no equity-efficiency tradeoff
  - Appealing combination of theory, structural, reduced-form, multiple fields

Introduction: What is Public Finance, What Will We Cover?

1. Taxation
   - **Equity**
   - **Efficiency**
   - Corporate taxes
   - Deficit finance
   - Behavioral
2. Externalities & public goods
   - **Cost-benefit analysis**
   - **Local public finance / fiscal federalism**
   - **Education, health, urban, environmental**
3. Social Insurance & redistribution
   - Unemployment insurance
   - Health insurance
   - Retirement
   - Disability insurance
   - Transfers (EITC, TANF, AFDC, etc.)

Motivation: Why Study Public Finance?

1. Important questions
   - Equity-efficiency tradeoff
   - Role of government
   - Market failures
   - Research can have tremendous social value
2. Appealing methods

Motivation: Why Study Public Finance?

1. Important questions
2. Appealing methods
   - Difference from other fields: emphasis on social welfare
   - Reduced-form meets structural estimates
     - Credible research designs for parameters of economic models
     - Goal is internal and external validity
     - "Sufficient statistics for welfare"
   - Natural links to many other fields
     - IO (especially for health, insurance, and energy/environment)
     - Labor (tax and expenditure programs)
     - Macro and theory (tax, insurance, fiscal policy)
     - Econometrics–less so; applied econometrics used widely
     - Development (tax, externalities, redistribution)
   - The course teaches what we know.
     - I'll also emphasize what we don't know, which are good areas for research

Overview

- Background on Public Finance/Environment
- This Course and Public Finance/Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

Course Requirements

- Each class we focus on 1-2 papers
- Criteria for choosing topics and papers:
  - Central questions in public finance/environmental/energy
  - Famous/classic papers
  - Demonstrate important tools
  - Examples of good research

Course Requirements

- 1 problem set
- Weekly written response papers
  - Due by 10pm night before class
  - Please email subject “ARE 261 reading response”
  - Make one constructive/critical point
  - Short
  - Do not summarize
- Referee report
- Participation
  - Questions, comments, discussion

Course Requirements

- Prerequisites
  - This course is at the level of a second-year course in the economics PhD sequence
  - May also be useful for graduate students in other programs
  - I strongly recommend you to have taken PhD-level microeconomics and econometrics
  - If you haven't and you want to take this class, talk to me

Course Overview

- Overview (today)
  - Big picture of empirical PF (structural, reduced-form, experiments, etc.)
- After today, many of the classes start with a paper focused on a central theory in the topic then follow with empirical applications of the theory

Overview

- Background on Public Finance/Environment
- This Course and Public Finance/Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

Sufficient Statistics for Welfare Analysis

Reduced form

- Also called policy evaluation, designed-based research, treatment effects, program evaluation
- Emphasis on transparency, causality, internal validity, omitted variables bias, selection bias, simultaneity bias
- Tends to be more data-driven empirical, simpler empirical methods

Structural

- Emphasis on utility-consistent foundations, external validity
- Identifies and estimates primitive economic parameters (utility, technology)
- More mathematical, complex empirical methods

"Third Way": combine the two approaches

Sufficient Statistics for Welfare Analysis

History

- Cowles Commission in 1950s developed simultaneous equations, structual models
- LaLonde (1986) helped initiate reduced-form research, led to skepticism about existing empirical results
- Program evaluation methods developed in 1990s (Angrist & Krueger handbook chapter, Angrist & Pischke book)

Sufficient Statistics for Welfare Analysis

A taxonomy of papers

1. Policy evaluation: show stylized facts or magnitudes
   - Possibly motivated by a model
   - Autor, Palmer, and Pathak; Walker
2. Policy evaluation: estimate quantities from a welfare formula
   - Deschenes, Greenstone, and Shapiro (2017)
3. Policy evaluation: estimate parameters allowing full welfare analysis
   - Sufficient statistics (Chetty 2009)

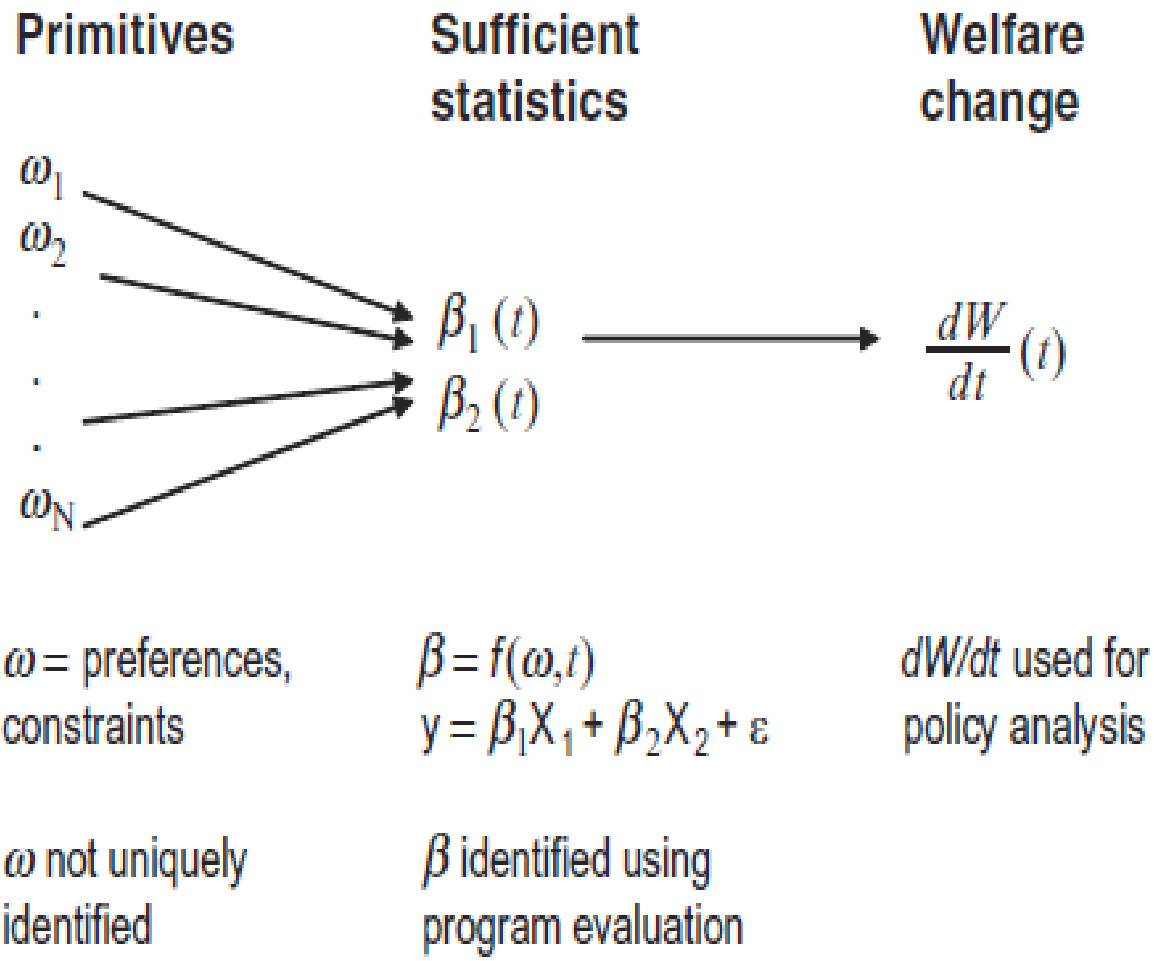
Sufficient Statistics for Welfare Analysis

A taxonomy of papers (continued)

4. Structural: identification incorporates policy evaluation tools
   - Ahlfeldt et al., Kremer et al.
5. Structural and policy evaluation: both estimate the same number. Check on internal validity.
   - Kline & Moretti (Forthcoming)
6. Structural model
   - Fowlie (2009)

Sufficient Statistics for Welfare Analysis

- In some cases, primitives not needed for welfare analysis of a specific question
- Main point of the paper



Overview

- Background on Public Finance/Environment
- This Course and Public Finance/Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

Harberger Triangles Example

Setup

- What is deadweight loss from taxing a good in a many-commodity general equilibrium economy?
  - Do we need to know cost functions, own-price and cross-price elasticities for all goods?
  - Does the calculation need to impose a market-clearing condition and see how prices adjust?
  - For marginal changes, no.
- Recall: what defines a competitive equilibrium (GE)?
  - Consumers maximize utility
  - Firms maximize profits
  - Markets clear

Harberger Triangles Example

Notation

- Z units of numeraire y is endowment, $p_y = 1$ normalization
- $(x_1, ..., x_J)$ other consumption goods firms produce from y, at price $p = (p_1, ..., p_J)$
- $c_j(x_j)$ firm cost function, $c(x)$ total cost of producing vector x
- t unit tax rate on good 1
- $p(t)$ market-clearing prices as function of taxes t

Harberger Triangles Example

Consumer utility maximization

$$\begin{aligned} & \max_{x,y} u(x_1, ..., x_J) + y \\ s.t. Z &= px + tx_1 + y \end{aligned}$$

Firm profit maximization

$$\max_x px - c(x)$$

Market clearing

$$x^D(p) = x^S(p)$$

Harberger Triangles Example

Planner's problem:

$$\begin{aligned} W(t) &= \{\max_x u(x) + Z - tx_1 - p(t)x\} + \{\max_x p(t)x - c(x)\} + tx_1 \\ &= \{\max_x u(x) + Z - tx_1 - c(x)\} + tx_1 \end{aligned}$$

In second equation, bracketed term is private surplus, tx_1 term is tax revenue

Important: consumers treat tax as parametric

Harberger Triangles Example

Goal: $\frac{dW}{dt}$, effect of tax on social welfare
Approaches:

- Structural: Estimate/calibrate $u(x)$ and $c(x)$.
  - Then compute $W(t)$ directly.
  - Tax on x_1 affects prices and quantities in all markets.
- Sufficient statistics: only need to know effect of tax on equilibrium quantity of taxed commodity, $\frac{dx_1(t)}{dt}$.

$$\begin{aligned} \frac{dW(t)}{dt} &= -x_1 + x_1 + t\frac{dx_1}{dt} \\ &= t\frac{dx_1(t)}{dt} \end{aligned}$$

Welfare change: $\Delta W = W(t_2) - W(t_1) = \int_{t_1}^{t_2} t\frac{dx_1}{dt}(t)dt$

Harberger Triangles Example

- Sufficient statistics uses two insights
  - In calculating $\frac{dW}{dt}$, envelope conditions let us ignore the $\frac{dx}{dt}$ term in curly brackets. Intuition: consumers and firms already chose x to maximize constrained welfare, so behavioral responses are second-order.
  - Prices are a transfer from consumers to firms so don't affect welfare, so we can ignore potential $\frac{dp}{dt}$ terms.
- So loss in social welfare equals gap between willingness to pay for x_1 and cost of x_1, which is a Harberger triangle.

Harberger Triangles Example

- Explanation #2. We could estimate

$$\frac{dx_1}{dt} = \frac{\partial x_1}{\partial p_1}\frac{\partial p_1}{dt} + \frac{\partial x_2}{\partial p_2}\frac{\partial p_2}{\partial t} + ... + \frac{\partial x_1}{\partial p_J}\frac{\partial p_J}{\partial t}$$

- Estimate $\frac{dx_1}{dt}$ directly rather than worrying about its components.

Harberger Triangles Example

- Limitations of the sufficient statistics approach
  - Ignores pre-existing distortions in other markets. For example, carbon taxes should account for inefficiency of pre-existing labor taxation. See Goulder and Williams JPE 2003.
  - Expression for marginal, not discrete changes in policy. Need strong functional-form assumptions for extrapolation.

Harberger Triangles Example

- What if utility differs across individuals?

$$u^i(x^i) + y$$

- Sufficient statitics result is same:

$$\frac{dW(t)}{dt} = t\frac{dx_1(t)}{dt}$$

- Intuition: effect of tax on aggregate demand describes marginal excess burden of tax. Doesn't matter which individuals bear the cost.
  - Unless we are worried about incidence or about extrapolation and external validity.
- What if individuals choose among discrete bundles of goods, possibly with logit errors?
- Same sufficient statistics formula:

$$\frac{dW(t)}{dt} = t\frac{dP_1(t)}{dt}$$

Harberger Triangles Example

- What if individuals choose among discrete bundles of goods, possibly with logit errors?
- Same sufficient statistics formula:

$$\frac{dW(t)}{dt} = t\frac{dP_1(t)}{dt}$$

Overview

- Background on Public Finance/Environment
- This Course and Public Finance/Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

General Setup: Sufficient Statistics

- Levy tax t to finance transfer $T(t)$
- Six steps:
  - Specify model structure
  - Express $\frac{dW}{dt}$ in terms of multipliers
  - Substitute multipliers by marginal utilities
  - Recover marginal utilities from observed choices
  - Empirical implementation
  - Model evaluation

General Setup: Specify model structure (I)

- $x = (x_1, ..., x_J)$ vector of choices for representative agent in private sector
- Tax t levied on good x_1, transfer $T(t)$ paid in units of x_J
- $\{G_1(x, t, T), ..., G_M(x, t, T)\}$ denote $M < J$ constraints
  - Budget constraints, restrictions on insurance or borrowing, hours constraints, etc.
- Agent's problem, treating t and T as given:

$$\max U(x) \text{ s.t. } G_1(x, t, T) = 0, ..., G_M(x, t, T) = 0$$

- Solution gives welfare as function of tax

$$W(t) = \max_x U(x) + \sum_{m=1}^{M} \lambda_m G_m(x, t, T)$$

General Setup: Express dW/dt in terms of multipliers (II)

- Using envelope conditions in private sector, differentiating W gives

$$\frac{dW}{dt} = \sum_{m=1}^{M} \lambda_m \left\{ \frac{\partial G_m}{\partial T} \frac{dT}{dt} + \frac{\partial G_m}{\partial t} \right\}$$

  λ_m is Lagrange multiplier for agent's constraint m
- $\frac{dT}{dt}$ is from government's budget constraint
- $\frac{\partial G_m}{\partial T}$ and $\frac{\partial G_m}{\partial t}$ calculated mechanically.
- Unknowns are multipliers λ_m. Represent marginal value of relaxing constraint: period-specific budget constraint, inter-temporal borrowing constaint, etc.

General Setup: Substitute Multipliers by Marginal Utilities (III)

- FOC of agent's problem give

$$u'(x_j) = -\sum_{m=1}^{M} \lambda_m \frac{\partial G_m}{\partial x_j}$$

  Inverting this system express multipliers in terms of marginal utilities.

- Generally use an assumption to simplify this inversion.

General Setup: Recover Marginal Utilities from Observed Choices (IV)

- Application-specific, but generally uses fact that marginal utilities are parts of FOC for choices.
  - In previous Harberger example, no income effects implies, $u'(x_J) = 1$.
  - Add FOC for x_1, which is $u'(x_1) = p_1 + t$
  - Into planner's problem gives

$$\begin{aligned} \frac{dW(t)}{dt} &= 1\left(x_1 + t\frac{dx_1}{dt}\right) - \frac{x_1}{p_1 + t}(p_1 + t) \\ &= t\frac{dx_1(t)}{dt} \end{aligned}$$

General Setup: Empirical Implementation (V)

- Suppose the sufficient statistic formula is

$$\frac{dW}{dt}(t) = f\left(\frac{\partial x_1}{\partial t}, \frac{\partial x_1}{\partial Z}, t\right)$$

  then construct empirical anlogues. Notes:

- Does the formula use partial or total derivatives? Varies by application.
  - Harberger measures total derivative $\frac{dx_1}{dt}$, which incorporates GE effects.
  - Reduced-form regressions may estimate the partial derivative $\frac{\partial x_1}{\partial t}$, holding prices in other markets fixed.
  - So reduced-form studies only useful for policy changes which don't affect prices in other markets.

General Setup: Empirical Implementation (V)

- Policy changes are never infinitesimal.
  - Ideally, we'd have nonparametric functions of t.
  - Then we measure $\frac{\partial x_1}{\partial t}(t)$ and $\frac{\partial x_1}{\partial Z}(t)$ and integrate between t_1 and t_2 to measure welfare gain ΔW. Similar to Heckman & Vytlacil MTE.
  - Often people estimate the LATE for a particular experiment:

$$\frac{\Delta x_1}{\Delta t} = \frac{x_1(t_2) - x_1(t_1)}{t_2 - t_1}$$

- Options with LATE
  - Bound welfare gain over this range: set slope between these points to lowest or highest possible values.
  - Approximate $x_1(t)$ to calculate dW/dt.

General Setup: Model Evaluation (VI)

- Test falsifiable predictions which are important for deriving sufficient-statistics formula.
  - Harberger triangles assume people treat taxes and prices identically.
  - This appears incorrect if taxes are not included in posted sales prices
- Identify at least one vector of structural parameters ω consistent with the estimated statistics

Overview

- Background on Public Finance/Environment
- This Course and Public Finance/Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

Application I: Taxation (Feldstein 1999)

- Many papers estimate efficiency cost of taxation as effect of taxes on hours worked
- But many channels for taxes to affect welfare. Training, effort, occupation, avoidance, etc.
- Structural approach: estimate each respose, aggregate
- Sufficient statistics: elasticity of taxable income with respect to taxes
  - Sufficient statistic for deadweight loss of taxation

Application I: Taxation (Feldstein 1999) Setup

- Government imposes tax rate t on reported taxable income
- Agent makes N labor supply choices $l_1, ..., l_N$
- Choice l_i has disutility $\psi_i(l_i)$, wage w_i
- Paying $g(e)$ lets an agent shelter e income from taxes
- Taxable income (TI):

$$TI = \sum_{i=1}^{N} w_i l_i - e$$

- Consumption is post-tax taxable income plus untaxed income:

$$x_N = (1 - t)\, TI + e$$

Application I: Taxation (Feldstein 1999) Sufficient Statistics Derivation

- Social welfare:

$$W(t) = \{(1-t)TI + e - g(e) - \sum_{i=1}^{N} \psi_i(l_i)\} + tTI$$

- Totally differentiating $W(t)$ gives

$$\frac{dW}{dt} = \frac{dTI}{dt} + \frac{de}{dt}(1 - g'(e)) - \sum_{i=1}^{N} \psi_i'(l_i)\frac{dl_i}{dt}$$

- FOC measure marginal utilities:

$$\begin{aligned} g'(e) &= t \\ \psi_i'(x_i) &= (1-t)w_i \end{aligned}$$

- Substituting into dW/dt gives Feldstein's main result:

$$\frac{dW}{dt} = t\frac{dTI}{dt}$$

Application I: Taxation (Feldstein 1999) Comments

- Intuition: optimization implies that marginal social cost of reducing earnings via each margin is equated at optimum.
  - So, irrelevant what mechanism causes decline in TI.
- Many studies estimate elasticity of taxable income
- Chetty (2008) undermines assumption $g'(e) = t$
  - Many avoidance behaviors are transfers, not genuine costs
  - Taxable income very sensitive to tax rates for rich, but efficiency cost may be smaller
- Structural approach would have estimated $g(e)$ so avoided this problem.

Application I: Taxation (Saez 2001)

- Mirrlees (1971) was foundational description of optimal tax theory
- Saez maps it to estimable elasticities using sufficient statistics
- Tax rate $T(z)$ at income z
- Net of tax income is $z - T(z)$

- Mirrlees: optimal tax in terms of primitives that enter FOCs
  - Doesn't clarify origins of $T(z)$
  - Simulation results depend on primitives

Application I: Taxation (Saez 2001), Mirrlees Model

- Workers choose labor supply. Worker problem:

$$\begin{aligned} u(c, l) &= c - \psi(l) \\ \text{s.t.} \quad c &= wl - T(wl) \end{aligned}$$

- Government chooses tax schedule $T(z)$ to maximize welfare:

$$W(T(z)) = \int_0^\infty \widetilde{G}(u(c(w, T), wl(w, T))dF(w)$$

Resource and IC constraints:

$$\begin{aligned} G_1(c, z, T) &= \int_0^\infty z(w, T)dF(w) - \int_0^\infty c(w, T)dF(w) - E = 0 \\ G_2(c, z, T) &= (1 - T'(z))w - \psi'(l(w)) = 0 \end{aligned}$$

Application I: Taxation (Saez 2001), Mirrlees Model

- Diamond and Saez obtain following formula for optimal tax $T(z)$:

$$\frac{T(z)}{1-T(z)} = \frac{1}{\varepsilon(z)zh(z)} \int_z^{\infty} (1-g(z'))h(z')dz'$$

- Elasticities $\varepsilon(z)$, density $h(z)$, marginal utility $g(z)$ at each point of income distribution determine optimal tax rate
- Saez uses to simulate $T(z)$

Overview

- Background on Public Finance/Environment
- This Course and Public Finance /Environment
- Sufficient Statistics: Overview
- Sufficient Statistics: Harberger Triangles Example
- Sufficient Statistics: General Setup
- Sufficient Statistics: Tax Applications
- Sufficient Statistics: Discussion

Sufficient Statistics Discussion

- Is a sufficient statistic for a specific question?
  - Depends on the model for that question.
  - New model, need new statistic.
  - In some settings, data can answer the question without filling in full model details. In most others, no.
- Many sufficient statistic formulas assume homogeneity in key dimensions, and other approaches or formulas needed when sufficient heterogeneity appears
  - e.g., Baily (1978) formula for optimal social insurance can be very misleading with heterogeneity in risk aversion
  - See Isaiah Andrews and Conrad Miller (2013 working paper)
- This discussion is really meta-research
  - You will do research, not meta-research
  - But envelope trick here is useful tool to remember.

Sufficient Statistics Discussion

- Reduced-form methods are well explained in a few settings
- Need to see many good examples to apply them well
  - Angrist and Pischke book
  - Angrist and Krueger
  - Differences in Differences
  - Differences in Differences in Differences
  - Regression Discontinuity
  - Instrumental Variables
  - Randomized Experiments
  - Imbens and Wooldridge's "What's New in Econometrics" 2007 lectures a great resource for extending basic methods.

Sufficient Statistics Discussion

- Structural methods–some general tools from microeconometrics
  - Criterion estimators (general method of moments, maximimum likelihood, simulation, etc.())
- Many others (value function iteration, contraction mapping)
- To some extent methods are more field- and topic-specific
- General suggestions:
  - Use the tools that are best suited to the problem at hand.
  - Learning tools before finding the problem is easier
  - Recognize: different fields and researchers have different views about what can be learned from various methods, including sufficient statistics.
  - For every paper you see here (and in general) ask: how/can I use this idea in my own research?
  - Also ask: why is this paper in the journal where it was published?