Class 8 - Techniques I - Regression

Agenda

- The logic of regression: what, why, when, how (40 minutes)
- Application paper discussion (20 minutes)
- *Break*
- Replication presentation (15 minutes)
- Skills corner - Class walkthrough in R (25 minutes)
- General discussion (15 minutes)

The logic of regression

Preamble

I will weave in the "core" papers as needed in my remarks later, but we will not discuss them in detail per se

WARNING: I am going deeper into the weeds this week, but we will come back up to a higher level soon

What is a regression?

The multiple linear regression **model** is used to study the relationship between a dependent variable (y) and one or more independent variables (X). The generic form of the linear regression model is (Greene 2012):

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where β is computed as follows:

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{y}) = \boldsymbol{\beta}$$

Let's figure out why...

What is a beta coefficient?

Small aside: Note that in a "simple" (bivariate) regression with just one predictor and an intercept

$$y = \beta_0 + \beta_1 x + \epsilon$$

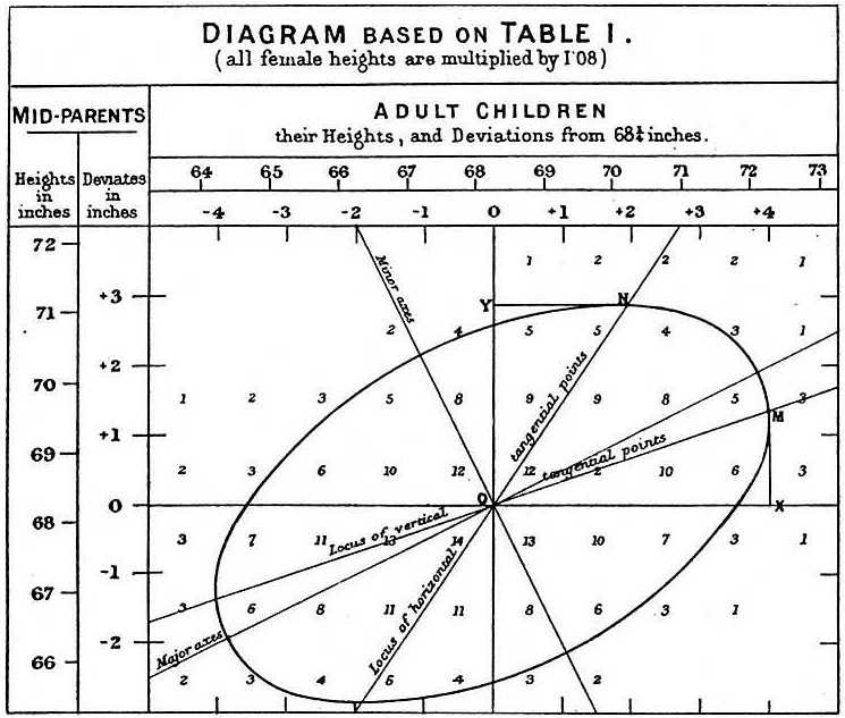
this funky formula is essentially:

$$Cov(x, y)/Var(x) = \beta_1$$

Note that in the multiple regression case, the same logic applies but to covariances and variances that have been "residualized" by accounting for all of the other X variables in the equation (Angrist and Pischke 2008)

What's with the name "regression"?

"Galton observed that extreme characteristics (e.g., height) in parents are not passed on completely to their offspring. Rather, the characteristics in the offspring **regress** toward a mediocre point (a point which has since been identified as the mean)" (Wikipedia)



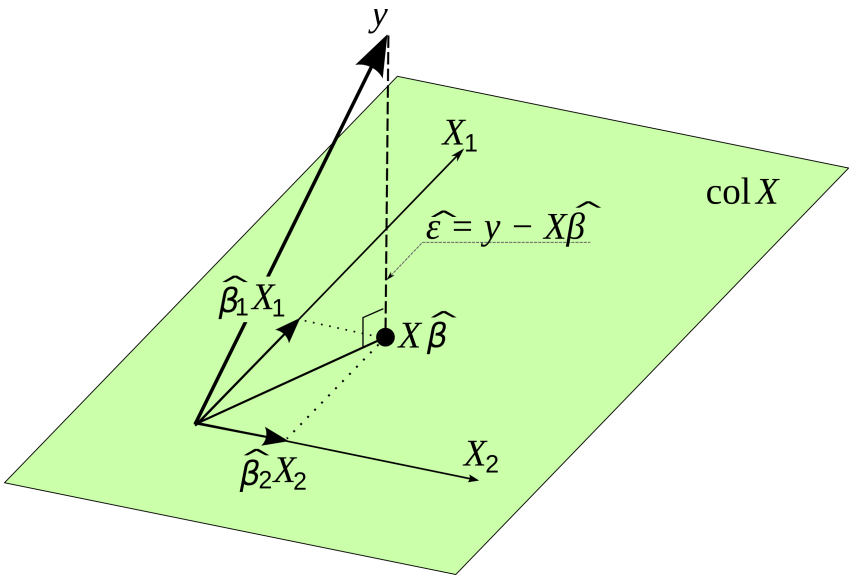
What do we usually mean when we say "regression"?

The term "regression" is often used synonomously with a particular technique: ordinary least squares (OLS).

- Essentially, OLS and its generalizations (e.g., WLS, GLS, discussed later) seek to minimize the distance between observed and predicted values
- It is constrained in this task in that only the data provided by X can be used, and that there are too many observations to simply interpolate

What is OLS doing?

- The data in X forms a (hyper)-plane, and each X variable is **weighted** by β, resulting in a prediction
- The discrepancy between this predicted value $X\beta$ and y is the error term ϵ, which by definition is unrelated to the X variables



A sample data generating process

Let's say we **know** that the relationship between X and Y is as follows:

$$y = (\beta_0 = 0) + (\beta_1 = 2)x + \epsilon$$

and that ϵ is randomly distributed with mean of 0 and a standard deviation of 3.

This is an example **data generating process**, what might such data look like?

A sample data generating process

```
# A tibble: 10 x 2
       x      y
   <dbl>  <dbl>
 1     1 -0.714
 2     2  7.28
 3     3  3.00
 4     4  6.05
 5     5 10.6
 6     6 10.2
 7     7 13.4
 8     1 -2.31
 9     2  3.78
10     3  7.14
```

OLS as a mathematical procedure

The population regression equation applies to each observation (thus the matrix form):

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{\epsilon}$$

and beta is **selected / defined** to minimize the sum of the squared errors (i.e., positive and negative errors don't cancel out):

$$min(\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}) = (\boldsymbol{y} - \boldsymbol{X}\beta)^T(\boldsymbol{y} - \boldsymbol{X}\beta)$$

In other words, we ask; what value of beta should we select to minimize the LHS?

OLS as a mathematical procedure

This minimization problem (with respect to beta) leads to the "normal equations" <- clink on the link for a full derivation.

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{y}) = \beta$$

Comments:

- What would happen the if error/disturbance term was always zero?
- Notice, this equation is agnostic to the form / distribution of the errors

OLS as an "estimator"

Remember, we don't have the whole population, we just have a sample. Thus, the **residuals** we get from an OLS estimation are different from the underlying **errors** in the population due to **sampling error** in our estimate of β

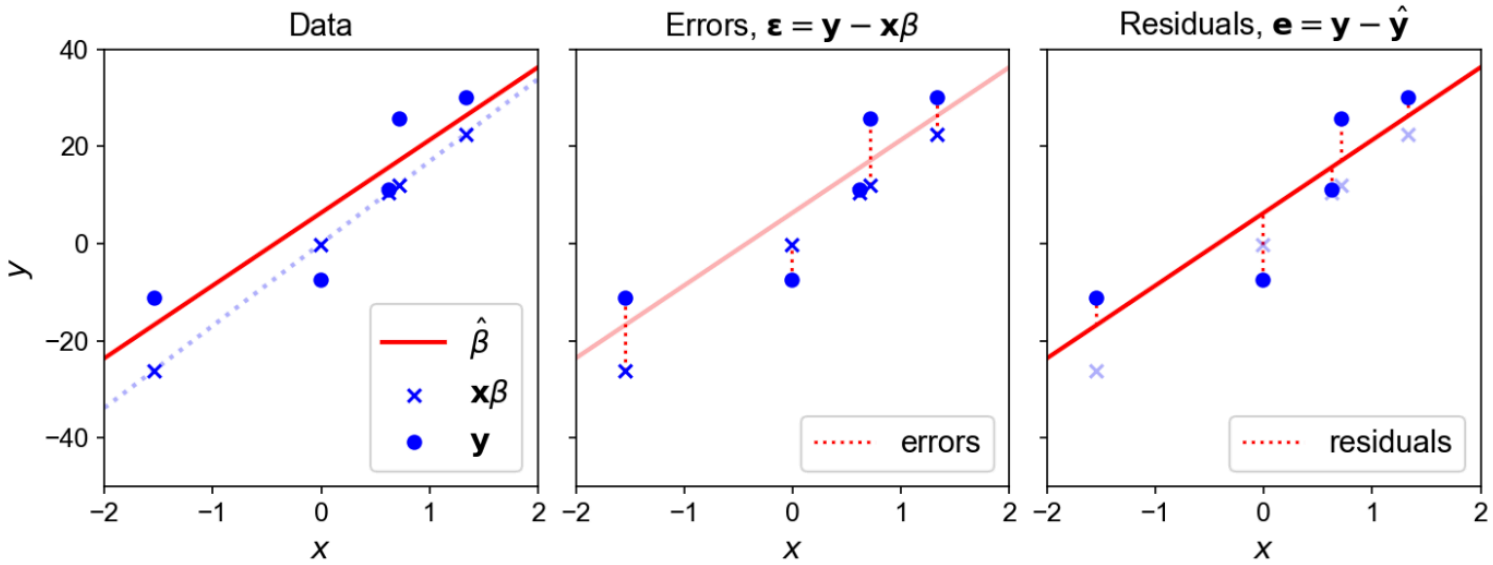


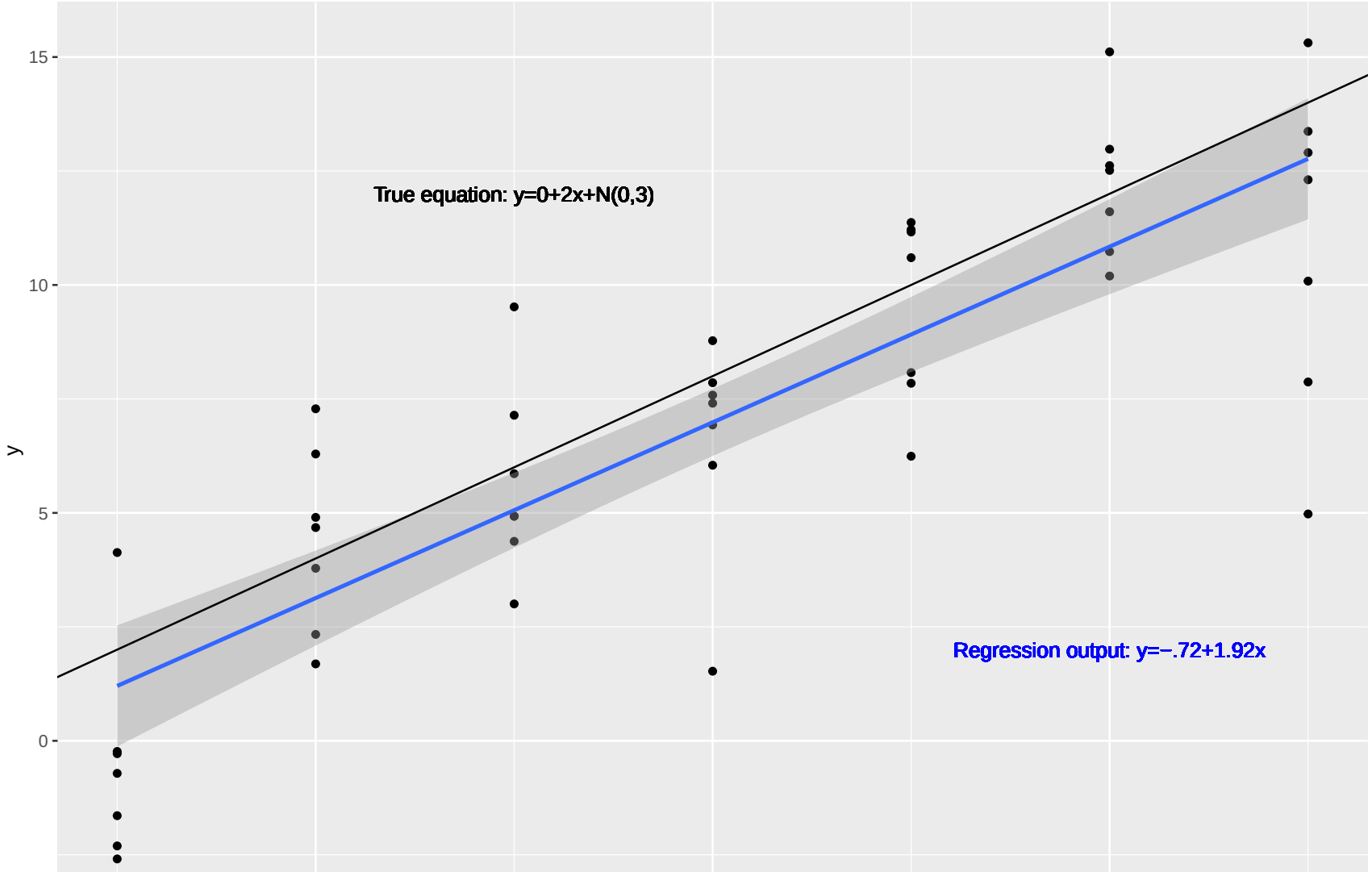
Figure 1. (Left) Ground-truth (un-observed) univariate data $\mathbf{x}\beta$, noisy observations $\mathbf{y} = \mathbf{x}\beta + \boldsymbol{\varepsilon}$, and estimated $\hat{\beta}$.

OLS as an "estimator"

Here is another example of this distinction with the data we generated earlier. Recall the "true" parameters:
$y = (\beta_0 = 0) + (\beta_1 = 2)x + \epsilon$

```
              Estimate Std. Error    t value     Pr(>|t|)
(Intercept) -0.7233802  0.8190561 -0.8831876 3.816302e-01
x            1.9272462  0.1831465 10.5229748 6.023326e-14
```

OLS - Estimation v. reality



Verifying our beta coefficient formula

Let's verify that our simple equation for $\hat{b}$ works out:

$$\hat{C}ov(x,y)/\hat{V}ar(x) = \hat{b}_1$$

```
[1] "cov(x,y) 7.8695885474943"

[1] "var(x) 4.08333333333333"

[1] "beta-hat 1.92724617489656"
```

What is a regression? - Prediction versus explanation

Notice a few things:

- OLS regression on a sample, by construction, minimizes the sum of the squared residuals
- This means that it seeks to maximize the amount explained by the regression to maximize the chances of correctly predicting the data **in the sample**
- By implication, it is NOT trying to find some true value of β such that we can make good **out of sample** predictions; issues with the sample will contaminate beta
- It also does NOT imply an causal interpretation or explanation

Why linear regression and OLS in particular?

OK, then why are regressions ubiquitous?

1. Approximates the conditional expectation function (Angrist and Pischke 2008)
2. The "best linear unbiased estimator" or BLUE, when the Gauss-Markov assumptions met (Kennedy 2008)
3. Equivalent to maximum likelihood estimator (MLE) and maximum a posteriori estimator (MAP) when errors normally distributed (Kennedy (2008), p. 43) and with a uniform prior

Let's consider each in turn.

"Approximates the conditional expectation function"

Recall that much of the information we want to extract from a distribution is summarized in its first two moments: mean and variance

The conditional expectation function [CEF] expresses the expected value (or mean) of a variable (Y) as a function of another variable (X)

The CEF is often not linear, but the OLS estimator is the best linear approximation to the CEF function (Angrist and Pischke 2008, Theorem 3.1.6)

"Approximates the conditional expectation function"

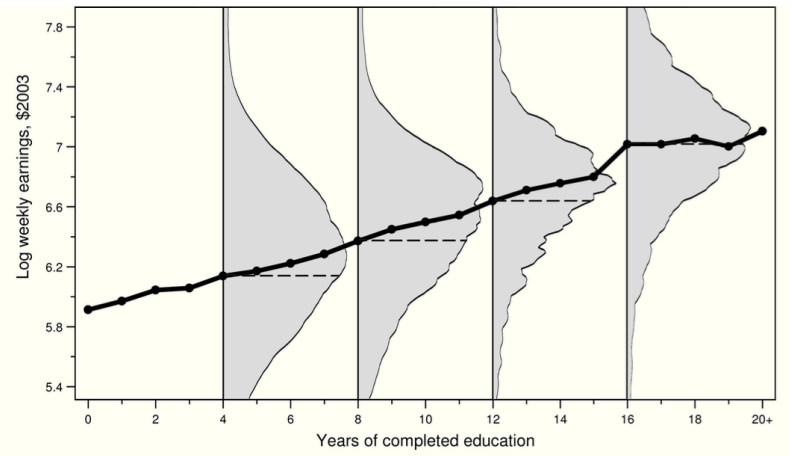
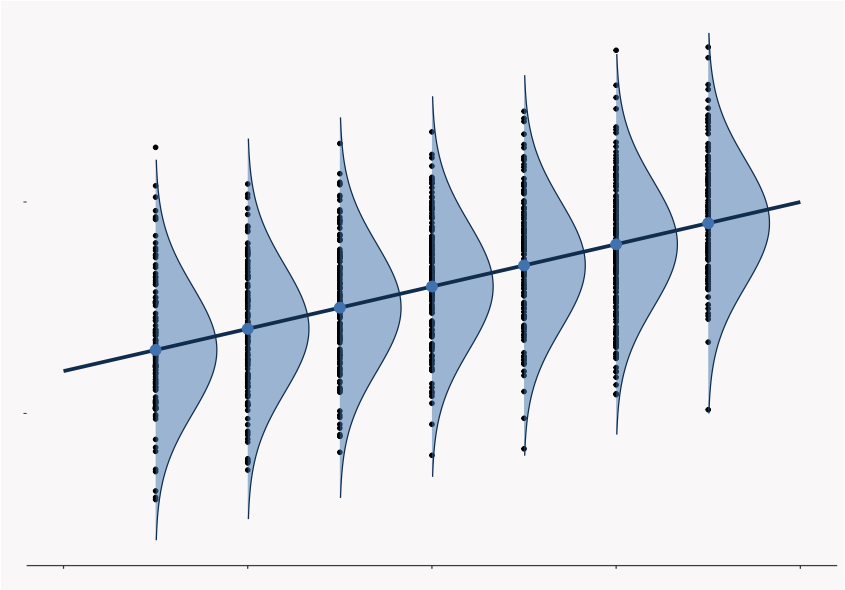


Figure 2: Angrist and Pischke (2008), Ch. 3

Note what this picture is showing you - it highlights the fact that Y has a distribution at each level of X, and what OLS can do is determine how that distribution "shifts" as you change X

"Approximates the conditional expectation function"

The "standard", unmodified form of OLS assumes that the distribution of Y remains the same across all levels of X (i.e., the distribution is homoskedastic or possesses "spherical errors"), and that the only difference is a shift in where the mean is placed



"Best linear unbiased estimator" (BLUE)

If the Gauss-Markov assumptions are satisfied, then OLS is the:

- **Best** (minimum variance among alternative estimators)
- **Linear**(the estimator is of the form $A(X)y$)
- **Unbiased** (the expected value is equal to the population parameter)
- **Estimator** (employs data from a sample to make inferences about the population)

Recent research seems to suggest that it might actually be the best unbiased estimator if these assumptions hold

"Equivalent to MLE and MAP under certain conditions"

There are other modeling frameworks to help make sense of data.

- Maximum likelihood estimation or MLE: What parameter values make the data most likely, GIVEN an assumed distribution?
- Maximum a posteriori (Bayesian) or MAP estimation: MLE that explicitly incorporates prior beliefs
- Generalized method of moments: A "generalization" of MLE that focuses on moments rather than distributions

Note that MLE and GMM are the primary alternatives in our field to OLS/WLS/GLS: they are used for SEM, HLM, panel data estimators, logit, probit, and other models

"Equivalent to MLE and MAP under certain conditions"

When we have a uniform prior (all parameter values within an interval are equally likely) and when the error term is normally distributed, OLS provides the same estimate as MLE and MAP.

This should be somewhat comforting, since it implies:

- The OLS estimate b is the most likely value for β, given the data and assuming the distribution is normal
- The OLS estimate is what we should update our beliefs to be if we did not have an informative prior

Aside: An example of how likelihood functions work

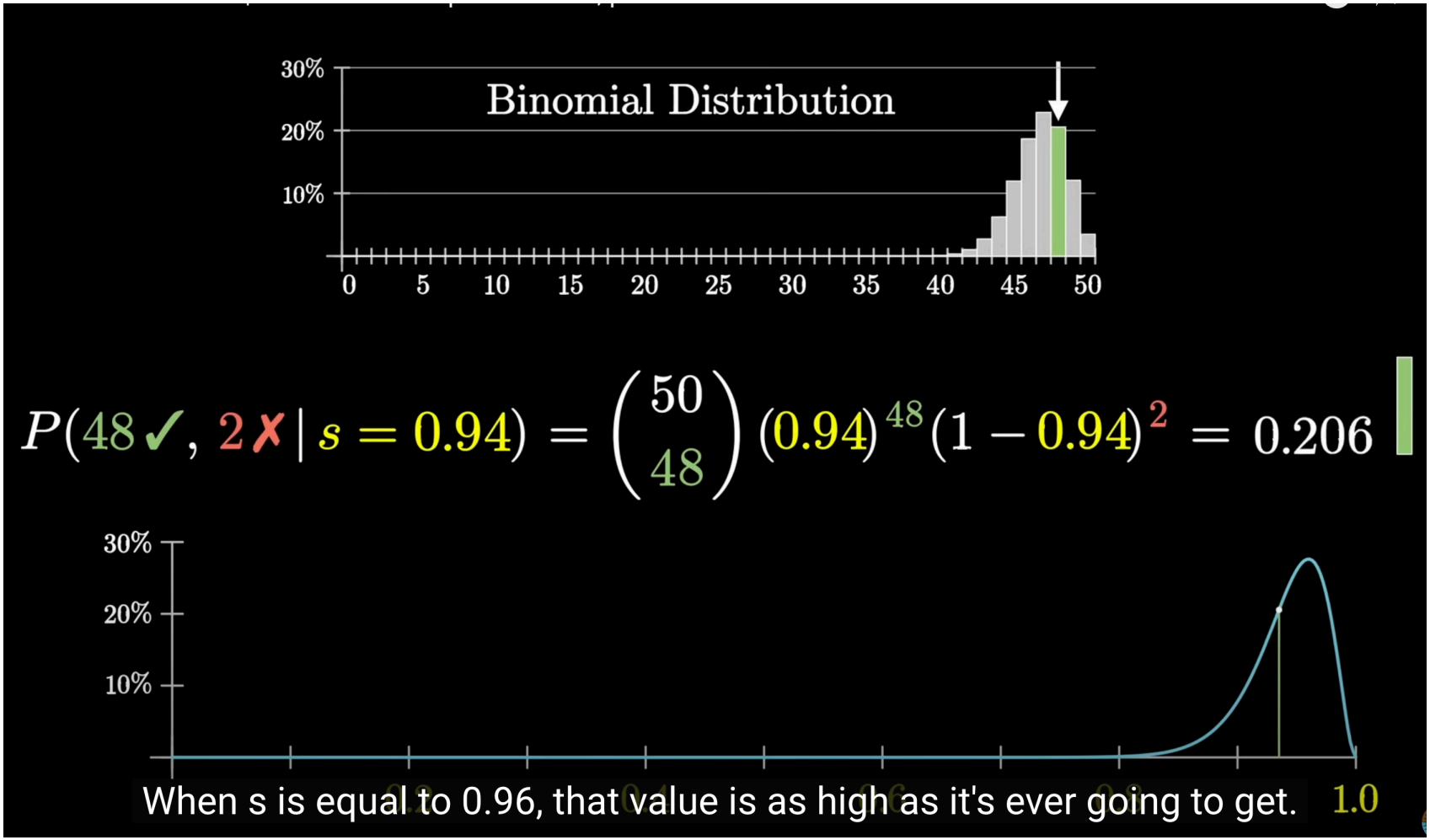


Figure 3: An example of a likelihood function

When is linear regression appropriate and / or optimal?

When the assumptions are satisfied!

And what are those assumptions? There are two "flavors":

- Fixed design
- Random design

I've tried to align the numbering so that they correspond, where possible.

When is linear regression appropriate and / or optimal?

Table 3.1 The assumptions of the CLR model.

Assumption	*Mathematical expression*		*Violations*	*Chapter in which discussed*
	Bivariate	*Multivariate*		
1. Dependent variable a linear function of a specific set of independent variables, plus a disturbance	$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$, $t = 1, \ldots, N$	$Y = X\beta + \varepsilon$	Wrong regressors Nonlinearity Changing parameters	6
2. Expected value of disturbance term is zero	$E\varepsilon_t = 0$, for all t	$E\varepsilon = 0$	Biased intercept	7
3. Disturbances have uniform variance and are uncorrelated	$E\varepsilon_t\varepsilon_r = 0, t \neq r$ $= \sigma^2, t = r$	$E\varepsilon\varepsilon' = \sigma^2 I$	Heteroskedasticity Autocorrelated errors	8
4. Observations on independent variables can be considered fixed in repeated samples	x_t fixed in repeated samples	X fixed in repeated samples	Errors in variables Autoregression Simultaneous equations	10 11
5. No exact linear relationships between independent variables and more observations than independent variables	$\sum_{t=1}^{N}(x_t - \bar{x})^2 \neq 0$	Rank of $X = K \leq N$	Perfect multicollinearity	12

The mathematical terminology is explained in the technical notes to this section. The notation is as follows: Y is a vector of observations on the dependent variable; X is a matrix of observations on the independent variables; ε is a vector of disturbances;

When is linear regression appropriate and / or optimal?

"Random design"

1. Linearity (model is correctly specified)
2.
3. Spherical errors $V[\boldsymbol{\epsilon}|\boldsymbol{X}] = \sigma^2 \boldsymbol{I}$
   - implies: Homoskedaticity $V[\epsilon|\boldsymbol{X}] = \sigma^2$
   - implies: No serial correlation $V[\epsilon_i \epsilon_j|\boldsymbol{X}] = 0; i \neq j$
4. Strict exogeneity $E[\epsilon|\boldsymbol{X}] = 0$
5. No multicollinearity (X is invertible)
6. Normality (optional) - allows us to make inferences about the sampling distribution of b

When is linear regression appropriate and / or optimal?

The Gauss-Markov theorem proves that, given these assumptions, the OLS estimator is the BLUE

- It becomes the best unbiased estimator if normality is assumed (it reaches the Cramer-Rao lower bound)
- The assumption of a spherical errors can be relaxed via Aitken's theorem, known as weighted (WLS) or generalized (GLS) least squares
- GLS often performs a transformation of the raw data to take into consideration the variance structure

Aside: alternative estimators when OLS/GLS not applicable

- Maximum likelihood estimation (MLE)
- Bayesian / maximum a posteriori (MAP) techniques
- Generalized method of moments (GMM) estimators
- Quantile regression
- Machine learning (e.g.,random forest models)
- LASSO (least absolute shrinkage and selection operation)
- Markov Chain Monte Carlo simulation and bootstrapping
- K-L divergence based metrics (e.g., Expectation-Maximization algorithms)

Performing a regression analysis

How to properly perform a regression analysis?

1. We determine the appropriateness of regression for our model and data from two angles:
   - Column (variable) perspective
   - Row (observation) perspective
2. Once we are satisfied that there are no issues (or that they have been addressed), run a "final, publication ready" regression.
3. Interpret the results using the tools of statistical inference (discussed later today).

The column (variable) perspective

Conditional mean structure

1. Do we have the right variables? (DAGs, theory)
2. Is the model in the correct functional form specified? (RESET,Chow tests)
3. Are the variables exogenous? (Sargen-Hansen, Durbin-Wu-Hausman)
4. Are your models well-conditioned? (VIF)

Variance structure

5. Is the mean structure correct? (see above)
6. Tests of heteroskedasticity and auto-correlation (e.g., Durbin-Watson, White tests)

The row (observation) perspective

Conditional mean structure

1. Are there outliers present? (Cook's distance)
2. Are there influential observations? (hat matrix)
3. Are certain observations missing or partially missing? (Censoring and truncation
4. Is the model dynamic in nature? (Arellano-Bond estimators)

Variance structure

5. Is there a nested structure to the data? (Level of analysis)

Deep dive: Do we have the right variables?

Note that we don't really have a "test" to apply here, but the assumption of proper specification is critical (including the right variables)

Why? Because if it is violated we don't only lose the "best" part of BLUE, we also lose the "unbiased" part, and then we are just "LE"

And furthermore, the estimator isn't even "consistent" -> i.e., gets closer to the true value as we increase the size of the sample, which is a "fall-back" position we can take if we don't have an unbiased estimator for smaller samples

The issue: Omitted variable bias

Omitted variable bias causes **ALL** of the beta estimates in a model to be biased in proportion the correlation between the regressors and the omitted variable left in the error term

This bias can be expressed in terms of "short" and "long" regression estimates, where short means a relevant variable b_2 has been omitted and the long is where it has been included (and it is assumed that the error term is truly uncorrelated in the case of the long equation):

In the simple case, the upshot is the short regression coefficient b_1 will be biased in the following manner:

$$b_{1short} = b_{1long} + b_{2omitted}(cov(b_1, b_2)/var(b_1))$$

The issue: Omitted variable bias

The problem: remember that OLS **assumes by construction** that the error term and the X variables are uncorrelated, so we can't rely on a quick test of their correlation to draw conclusions. You need to **know from theory or experience** that you are missing the variable and then can run the test of including it to see if it changes the results (assuming it is something you can measure!)

What can you do if this variable is unobservable? [See our endogeneity day!]

The flip side: The "illusion of statistical control"

By mapping practices to these purposes, we demonstrate why current CV practice struggles to accomplish any of them effectively while potentially reducing the interpretability of results. Empirically, we examine correlations between CVs and independent variables (IVs)—relationships that receive little consideration in standard CV practice—demonstrating these relationships can and do influence the magnitude and sign of regression coefficients, sometimes dramatically—just not very often. In fact, the CVs in the studies we reviewed most often have little if any impact on research find- ings or interpretations, creating the illusion of statistical control when little control actually occurs. (Carlson and Wu 2012, 415)

What if we fail to meet one or more of the assumptions?

An econometrics textbook can be characterized as a catalog of which estimators are most desirable in what estimating situations. Thus, a researcher facing a particular estimating problem simply turns to the catalog to determine which estimator is most appropriate for him or her to employ in that situation. The purpose of this chapter is to explain how this catalog is structured. (Kennedy 2008, 40)

What if we fail to meet one or more of the assumptions?

But note that:

> *If more than one of the CLR model assumptions is violated at the same time, econometricians often find themselves in trouble because their catalogs usually tell them what to do if only one of the CLR model assumptions is violated. Much recent econometric research examines situations in which two assumptions of the CLR model are violated simultaneously. These situations will be discussed when appropriate. (Kennedy 2008, 44)*

Statistical inference

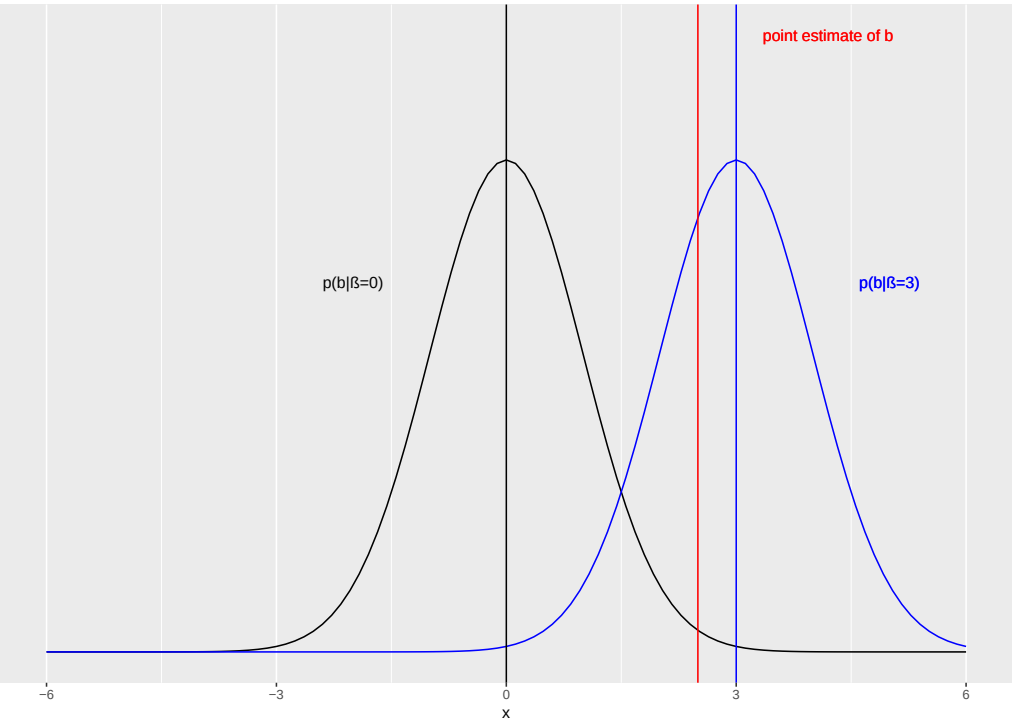
Are our results "statistically significant"?

Once we have an estimate for β (known as $\hat{b}$), here are two approaches to statistical inference (Kennedy 2008):

- Explicitly assume that the error term is distributed normally (t-tests and F-tests are appropriate straight away, even for small samples)
- Rely on the central limit theorem and large-N asymptotics (beta is an expected value, so its distribution will convergence to a normal distribution - see this video to understand why)

The distribution of $\widehat{b}$

When making an inference, we ask where our point estimate (what we got from the regression) sits within the distribution of the estimator conditional on some assumed true value of β



Determining appropriate standard errors

Questions about the variance structure also need to be settled:

- Are errors spherical?
- If not, do you want to incorporate the error structure information into the estimates (using GLS) or use OLS and "robust" statistics"?
- Two schools of thought: Feasible GLS has no guarantees of improvement, but "Simply computing a robust covariance matrix for an otherwise inconsistent estimator does not give it redemption" (Greene 2012, 692)

Another word about "robust standard errors"

Per Angrist and Pischke (2008), pp. 40-48 - robust standard errors are computed using the method of moments, and justified asymptotically by the law of large numbers / central limit theorem (regardless of the underlying distribution of the estimator) Thus, the results are **consistent** regardless of the covariance structure but you need to worry about finite-sample bias

Statistical inferences you can draw

The simplest test:

- Is $\beta \neq 0$?: Use t-tests

Intuition: We assume a null distribution that $\beta = 0$, and then see where our estimated value falls within that distribution. The more extreme it is, the more likely the assumption that $\beta = 0$ is false

Statistical inferences you can draw

A more powerful set of tests: linear hypotheses

- Is $\beta_1 = \beta_2$?
- Is $\beta_1 = -\beta_2 + 2\beta_3$?
- Is $\beta_1 = \beta_2 = \beta_3 = 0$?

A restriction matrix is constructed to run an F-test (Wald test), Likelihood Ratio (LR) or LaGrange Multiplier (LM) test (Kennedy 2008, ch. 4)

Intuition: Unconstrained optimization (remember we are minimizing least squares) is easier than constrained. We ask whether there is a significant difference in model fit after imposing the restriction: If yes, then the restriction is probably false.

Examples:

"Linear hypothesis tests fail to reject the assertion that growth and reductions in either team size ($\chi^2(1) = 0.39$, $p > 0.10$) or within-industry experiences ($\chi^2(1) = 0.17$, $p > 0.10$) have equal and opposite effects" (Fox, Simsek, and Heavey 2023, 17)

"A Wald test indicates that the hypothesis that both variables are simultaneously zero is rejected ($\chi^2(2) = 7.50$, $p = .02$)" (Fox, Simsek, and Heavey 2022, 18, FN)

Aside: Statistical inference and p-hacking

- "Searching for asterisks" is all about demonstrating that a beta coefficient is "statistically significant" and differs from 0.
- It does not consider economic / practical effect size
- There are many tools in the toolbox to "manage" statistical significance
- More useful things to focus on are effect size, confidence intervals, and robustness to different specifications of mean and variance structures

Statistical inferences and p-hacking

On the flip side, being careful when performing your analyses can stop you from making faulty inferences - even if it is painful in the short run.

> *T-stat looks too good. Use standard errors- significance gone. -Keisuke Hirano (Angrist and Pischke 2008, 6)*

The flip side of significance: Statistical power

Don't forget to give yourself a fighting chance to find the effect you are looking for!

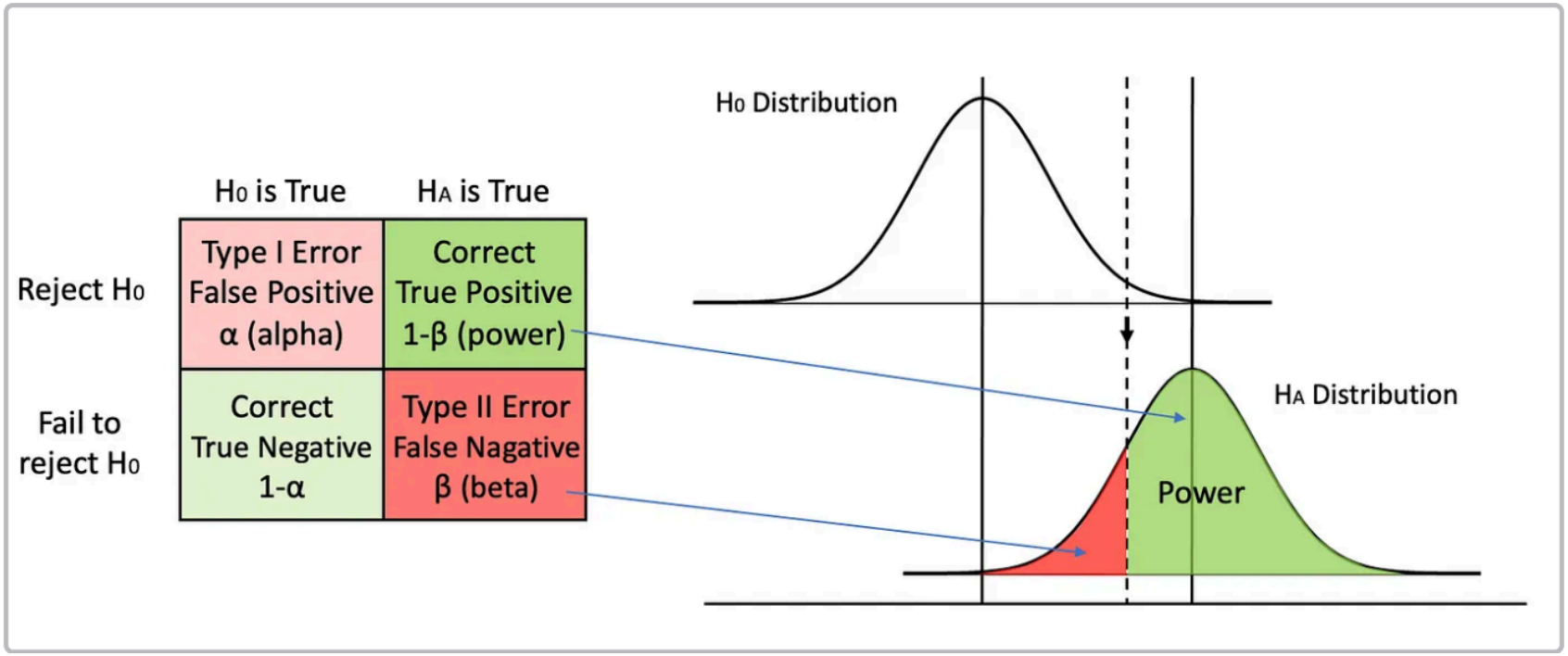


Figure 5: Source: Yao 2021

Applications

Application readings

Let's level-set people's familiarity with these pieces.

- Katila, R., & Ahuja, G. 2002. Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. Academy of Management Journal, 45(6), 1183-1194.
- Simsek, Z., Fox, B., & Heavey, C. 2021. Systematicity in Organizational Research Literature Reviews: A Framework and Assessment. Organizational Research Methods, 109442812110086.

Katila and Ahuja (2002)

- What was this paper about?
- What were the findings?
- What was the method?
- What makes sense? What was confusing?

© Academy of Management Journal
2002, Vol. 45, No. 6, 1183–1194.

SOMETHING OLD, SOMETHING NEW: A LONGITUDINAL STUDY OF SEARCH BEHAVIOR AND NEW PRODUCT INTRODUCTION

RIITTA KATILA
Stanford University

GAUTAM AHUJA
University of Michigan

We examine how firms search, or solve problems, to create new products. According to organizational learning research, firms position themselves in a unidimensional search space that spans a spectrum from local to distant search. Our findings in the global robotics industry suggest that firms' search efforts actually vary across two distinct dimensions: search depth, or how frequently the firm reuses its existing knowledge, and search scope, or how widely the firm explores new knowledge.

In this study, we examined how firms *search*, or solve problems (Nelson & Winter, 1982), to create new products. The ability to create new products is an important component of firm innovative capabilities. New products are a central mechanism whereby organizations diversify, adapt, and reinvent themselves in changing market and technical conditions (Schoonhoven, Eisenhardt, & Lyman, 1990). Research has also demonstrated how new products improve the market share, market value, and survival of firms (Banbury & Mitchell, 1995; Chaney & Devinney, 1992). Yet, despite the attractiveness, firms find it difficult to create new products. Here, we explain a firm's performance in creating products as a function of its search behavior.

Organizational learning researchers have sometimes argued that in their search for solutions to problems, firms position themselves in a unidimensional search space that spans the spectrum from exploitation to exploration (March, 1991). We suggest that firms' search, or problem-solving efforts, actually vary on two distinct dimensions rather than one. Firms can vary in their degree of use and reuse of their existing knowledge, just as they can vary in their exploration of new knowledge. We call the first dimension, which describes how deeply a firm reuses its existing knowledge, *search depth*. We call the second dimension, which describes how widely a firm explores new knowledge, *search scope*. In the sections that follow, we develop and apply this framework to the context of new products and argue that a firm's ability to create new products is determined by the independent and interactive effects of search depth and search scope.

We thank Avi Fiegenbaum, Jim Fredrickson, Andy Henderson, George Huber, David Jemison, Paul Mang, and Ken G. Smith for comments and discussions on the earlier versions of this work. We are also grateful to a number of experts and researchers in the robotics industry for their time, insights, and suggestions. We especially thank Risto Miikkulainen for assistance in writing the programs that were used to combine and analyze the patent and product data. Financial support from the Abell-Hanger Foundation, the Center for Customer Insight at the University of Texas at Austin, and the National Science Foundation (Grant #0115147) is acknowledged. An earlier version of this paper was selected as a

CONCEPTUAL BACKGROUND

The core technical and user service features of a product are customarily called a product's design (Saviotti & Metcalfe, 1984). In this study, a *new product introduction* was defined as any change in a product's design. New products represent the potential commercial value of a firm's R&D activities; most innovations do not influence firm performance until they are introduced to the market. A construct of product introductions also complements other, more intermediate proxies for firm innovation, such as knowledge, R&D investment, and scientific publications. Yet relatively few longitudinal studies have explored the determinants of new product introductions.

Search in organizations is one part of the organizational learning process through which firms at-

Simsek Fox and Heavey (2021)

- What was this paper about?
- What were the findings?
- What was the method?
- What makes sense? What was confusing?

Feature Topic: Rigorous and Impactful Literature Reviews

Organizational Research Methods
2023, Vol. 26(2) 292–321
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1094428121008652
journals.sagepub.com/home/orm
SAGE

Systematicity in Organizational Research Literature Reviews: A Framework and Assessment

Zeki Simsek[1], Brian Fox[2], and Ciaran Heavey[3]

Abstract

In this study, we first develop a framework that presents systematicity as an encompassing orientation toward the application of explicit methods in the practice of literature reviews, informed by the principles of transparency, coverage, saturation, connectedness, universalism, and coherence. We then supplement that conceptual development with empirical insights into the reported practices of systematicity in a sample of 165 published reviews across three journals in organizational research. We finally trace implications for the future conduct of literature reviews, including the potential perils of systematicity without mindfulness.

Keywords

literature reviews, systematic reviews, systematic literature reviews, systematicity

The growing volume and diversity of studies, proliferation of searchable databases, and rapid diffusion of research have contributed to the increased prevalence of literature reviews (Briner et al., 2009; Rousseau et al., 2008). Concurrently, the role and purpose of literature reviews in organizational research have evolved considerably over the past quarter of a century. Traditionally employed to summarize the state of the science, literature reviews are increasingly utilized to integrate, refine, and elaborate theory (Baumeister & Leary, 1997; Grant & Booth, 2009; Leidner, 2018; Post et al., 2020; Rowe, 2014; Sandelowski & Barroso, 2003; Siddaway et al., 2019; Tranfield et al., 2003; Xiao & Watson, 2019). As scientific disciplines become increasingly diversified, literature reviews also provide an antidote to proliferation and balkanization by bridging diverse theories and approaches that address shared questions.

Perhaps not surprisingly, researchers, practitioners, and policymakers have discussed the need for and importance of *systematicity* to strengthen the methodological rigor of literature reviews. Quite

[1]Department of Management, College of Business, Clemson University, Clemson, SC, USA

Break

COFFEE BREAK

Replication Presentation

- Replication: Simsek, Z., Fox, B., & Heavey, C. 2021. Systematicity in Organizational Research Literature Reviews: A Framework and Assessment. Organizational Research Methods, 109442812110086.

Skills corner

Class walkthrough in R

- Our "skills corner" for the rest of the semester is embedded in two parts of the class:
  - The "how to" discussion in the first hour, talking about the general process
  - The "workshop" now where we apply these ideas in the context of R
  - Let's fire up R and see how to execute

General discussion

Preparation for next class

Next class

Concept check 2 will available this evening - please complete it before next class.

Next class

Techniques II: Moderation

1. Dawson, J. F. 2014. Moderation in Management Research: What, Why, When, and How. Journal of Business and Psychology, 29(1), 1-19.
2. Hitt, M. A., Beamish, P. W., Jackson, S. E., & Mathieu, J. E. 2007. Building Theoretical and Empirical Bridges Across Levels: Multilevel Research in Management. Academy of Management Journal, 50(6), 1385-1399.

Next class

Techniques II: Moderation

Applications:

3. Replication: Heavey, C., Simsek, Z., & Fox, B. C. 2015. Managerial Social Networks and Ambidexterity of SMEs: The Moderating Role of a Proactive Commitment to Innovation. Human Resource Management, 54(S1).
4. Wolfson, M. A., & Mathieu, J. E. 2018. Sprinting to the finish: Toward a theory of Human Capital Resource Complementarity. J Appl Psychol, 103(11), 1165-1180.

References

Angrist, J. D., and J. S Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Carlson, Kevin D., and Jinpei Wu. 2012. "The Illusion of Statistical Control." *Organizational Research Methods* 15 (3): 413–35.

Fox, Brian C., Zeki Simsek, and Ciaran Heavey. 2022. "Top Management Team Experiential Variety, Competitive Repertoires, and Firm Performance: Examining the Law of Requisite Variety in the 3D Printing Industry (1986–2017)." *Academy of Management Journal* 65 (2): 545–76.

———. 2023. "Venture Team Membership Dynamics and New Venture Innovation." *Strategic Entrepreneurship Journal*.