Endogeneity and the Importance of Quasi-Experiments for Causal Inference in Accounting Research

Vinicius Gomes Martins
https://orcid.org/0000-0001-7401-9570

Giuseppe Trevisan
https://orcid.org/0000-0001-5165-0597

1. Motivation

Research in the accounting field underwent a significant paradigm shift, notably from the 1960s onwards, driven by studies by Ball and Brown (1968) and Beaver (1968). These pioneering studies introduced the positivist perspective into Accounting, which began to emphasize the empirical and quantitative approach. With the development of economic theory and finance, as well as the advancement of information technology and the increasing availability of financial data, the adoption of an approach based on quantitative evidence has intensified in accounting research.

This structural shock profoundly impacted how researchers conduct their studies and the demands imposed by the leading Accounting journals, as empirical studies with a quantitative focus began to be increasingly valued. Consequently, studies without a quantitative design are virtually impossible to be published in many periodicals. Thus, there has been a significant increase in the production of studies adopting econometric modeling to identify relationships between variables, predict behaviors, and test theories seeking to explain phenomena involving Accounting.

As a result of this process, the most significant research problems in mainstream accounting research involve the use of observational data and the search for evidence of causal effects (Gow et al., 2016). For example, it is common to come across studies seeking to analyze the "impact," "effect," or "influence" of a variable X on a variable Y. All of these works implicitly demand causal inference. However, the designs adopted by many of these studies have generated an increasing number of questions about the effectiveness of the empirical strategy and raised doubts about whether the results can actually reveal causal effects or are restricted to presenting limited descriptive correlations.

The main problem in interpreting a causal relationship associated with these studies lies in Endogeneity, something prevalent in studies using accounting data. Overcoming the Endogeneity problem is a recurring challenge since accounting variables are often subject to mutual influences and factors difficult to control. These aspects may lead to spurious correlations or biases in the estimated parameters, compromising the ability to identify causality and, therefore, reflecting on the results' reliability.

The challenges associated with studies seeking to capture causal relationships are increasingly evident. Fortunately, researchers have addressed these issues and improved the empirical strategies to identify cause-and-effect relationships. It is worth highlighting the contributions of the 2021 Nobel Prize winners in Economics, David Card, Joshua D. Angrist, and Guido W. Imbens, who are recognized for using quasi-experimental techniques for causal inference, especially in fields such as labor and education economics. These economists' innovative approaches have gone beyond Economics and influenced other research fields, including Accounting.

Based on the previous discussion, this editorial presents some reflections on causal research in Accounting, emphasizing the Endogeneity problem. We propose an introductory discussion on the importance of quasi-experiments, especially about feasible techniques that can be adopted by many research projects in the field. Among these techniques, we highlight the Difference-in-Difference design, the use of Instrumental Variables, and the Regression Discontinuity design. Thus, our purpose is to encourage the adoption of more rigorous identification methods to sensitize researchers about the quality of research in Accounting, paving the way for a more solid and innovative accounting research environment.

2. The Endogeneity Problem

Ultimately, what is Endogeneity, and why should it receive attention from researchers when designing studies intended to interpret a causal relationship? Objectively, Endogeneity occurs when the explanatory variable of a regression model is correlated with the error term, resulting from the omission of relevant variables, reverse causality, and/or measurement error in the regressor. Therefore, it represents a violation of the Exogeneity hypothesis, which is essential for deriving unbiased estimators using the Ordinary Least Squares (OLS) method. Hence, researchers must be alert to this matter when interpreting the estimate of the parameter of interest since any bias in the estimator compromises the causal inference between the regressor and the dependent variable of the econometric model.

Thus, we bring a didactic example to better illustrate the problem caused by Endogeneity. For brevity, we focus the discussion on the first two types of Endogeneity previously mentioned. Suppose you are interested in measuring the effect of belonging to the B3 Corporate Sustainability Index (ISE) portfolio on the performance of publicly traded companies. Let the performance measure of company i be Y_i, the binary variable D be an indicator of the company's adherence to the ISE (=1) or (=0) otherwise and, finally, let X_i ($X1_i$, $X2_i$, ..., X_{ki}) be a vector that contains "k" observable characteristics. Initially, we consider a context in which data are observed through a cross-section. The empirical model can be expressed as:

$$Y_i = \beta_0 + \beta_1 D_i + X'_i \gamma + \varepsilon_i, \qquad (1)$$

where ε_i represents the regression error term, and β_1 is the parameter of interest, which should measure the "effect" ("impact" or "influence") of ISE on business performance. However, a priori there is no reason to believe that "β_1" represents the average effect of interest. This is because, without understanding the factors that determine the company's membership to the ISE, nothing ensures that this parameter will be estimated without bias. In this sense, some points in Equation 1 should be highlighted.

First, the k variables contained in vector X_i will not contribute to minimizing bias in the estimation of β_1 if they are not **predetermined factors concerning adherence to the ISE**. It occurs because the characteristics of this vector may reflect the behavior of companies in response to joining the ISE[1] and even play the role of dependent variables in the model. Angrist and Pischke (2009) highlight this problem and call these variables "bad controls." Hence, the variables in X_i may not be good candidates for control variables. Second, due to the simple fact that the econometric model represents a simplification of reality and reflects the researcher's lack of knowledge about several phenomena in its stochastic term ε_i, **other unobservable** characteristics(or confounding factors) that are relevant and correlated with adherence to the ISE may exist. Consequently, omitting variables determining a company's performance and likelihood of joining the ISE will imply $\hat{\beta}_1$ bias. Therefore, the companies' performance variation may be partially or entirely caused by a factor not observed by the researcher and, fallaciously, be attributed to the companies' adherence to the ISE. The two points discussed here reflect the problem of Endogeneity well. It is also worth highlighting that the magnitude of such bias can be so significant to the point of affecting the estimate's direction, compromising both the interpretation of the correlation's "impact" and "sign."

---

1   For example, considering the actual eligibility criteria for the ISE portfolio, it is relatively simple to argue that joining the ISE is endogenous to a company's size since larger companies tend to self-select to participate in the portfolio and present better performance. Additionally, the size of the participating companies may be impacted as a reflection of ISE membership. Thus, using the natural logarithm of total assets (a traditional proxy for company size) in the empirical model as a control variable can be quite problematic. Ahern and Dittmar (2012) propose a similar discussion by estimating the effect of changes in the Board of Directors on the valuation of corporations. Larcker et al. (2007) propose another interesting example and discuss endogeneity in using the leverage variable in their econometric specification when analyzing the association between corporate governance and organizational performance.

Another source of estimating bias concerns the reverse causality imbued in the relationship of interest. In our context, both arguments that "joining the ISE can affect a company's performance" and "the level of performance may determine a company joining the ISE" make sense. We must agree that the chronological issue is crucial to understanding the causal relationship since the cause precedes the effect. However, we note that exploiting longitudinal information in the data is insufficient to overcome the problem of reverse causality — which, from an alternative perspective, can be seen as an omitted variable problem — and obtain the desired causal interpretation.

To motivate this issue, we will include a temporal dimension in the previously declared variables of our example, assuming annual periodicity ($t = \{0, 1, 2, ..., T\}$). When considering longitudinal data, the most convenient starting econometric specification for estimating the relationship of interest would be a fixed effects model[2]:

$$Y_{it} = \beta_0 + \beta_1 D_{it} + X'_{it}\ \gamma + \alpha_i + \alpha_t + u_{it}, \qquad (2)$$

where u_{it} represents the regression error, and the terms α_i and α_t are, respectively, company and year fixed effects. Including the first fixed effect is essential to control for idiosyncratic factors of companies that are invariant over time and cannot be observed by the researcher, e.g., the year of incorporation and sector of activity. The year-fixed effect absorbs temporal shocks distributed homogeneously among companies (e.g., inflation, seasonality, interest rates, and macroeconomic shocks, among many others). Using the longitudinal data structure of the data does not solve the problem of reverse causality because there may be temporal trends in the dependent variable that precede the company's participation in the ISE. In other words, dynamically, the level of performance may be responsible for the company's choice to participate in the ISE portfolio. On the other hand, note that the opposite is also an empirically susceptible situation. The central issue is that, in this context, the researcher cannot observe either the first or the second situation in the data, allowing a feedback relationship between the variables.

Furthermore, even when using more robust econometric modeling, as in Equation 2, the problem of identifying the causal effect remains due to the same problems previously discussed. It is worth noting that the fact that joining the ISE constitutes a company decision, other unobservable time-varying factors may be associated with the company's choice to participate in the portfolio and which directly affect its performance. For example, the company may decide to participate in the portfolio motivated by the adherence behavior of its peers in the same sector, which may reflect on its performance. All the points discussed in this section regarding the difficulty in estimating the causal effect translate the phenomenon known as **selection into unobservable**, i.e., selection bias caused by the existence of factors the researcher cannot control. In short, it is impossible to infer causality between performance and ISE if the self-selection behavior of companies in the ISE portfolio remains unresolved.

---

2   It is worth noting that since the decision to participate in the ISE portfolio comes from companies, there is no reason to believe that a random effects model (or even a pooled OLS) is employable in this context. It happens because of the individual's unobserved heterogeneity, "" is not independent of the variable , reflecting this relationship's endogeneity. Furthermore, Hausman will be uninformative regarding the decision to use one model over another due to its sensitivity regarding the inclusion of regressors. In fact, Wooldridge (2015) notes that "*if we think the unobserved effect α_i is correlated with any explanatory variables, we should use first differencing or fixed effects*".

The approach thus far is very provocative and somewhat discouraging concerning the consequences Endogeneity brings to causal interpretation. Thus, the question becomes pertinent: Are there ways to get around this problem? Fortunately, yes. However, it should be noted that the solutions are not so trivial and are not always feasible. In this sense, the perception of implementation opportunities requires considerable insight from the researcher regarding the institutional context of what is being investigated. Before we present the solutions, it is helpful to define some key terms.

Let us consider a **factual** situation in which a group of individuals receives an intervention of any nature (**treatment**) and a **counterfactual** situation in which that same group does not receive the intervention. The purpose is to estimate the average treatment effect on a given **impact indicator** (dependent variable), measured after the intervention. The average impact of interest could then be obtained by the difference in the impact indicator averages between the factual and counterfactual worlds. However, note that while the first situation is tangible, the counterfactual is not (perhaps, in a parallel universe, it is). This impossibility of observing the counterfactual world creates what is known as the Fundamental Problem of Causal Inference[3]. The researcher's objective is, then, to find a group of individuals in the real world that mimics the counterfactual situation of the treated group (called the **control group**). In other words, it is necessary to identify individuals who are comparable to those selected to receive treatment. An important message is that not all observational units that make up the untreated group constitute a good control group in the impact analysis.

Experiments are the "gold standard" to determine causal inference. Experiments are interventions conducted by researchers or policymakers in which the treatment condition is randomly manipulated[4]. Randomization intends to ensure that the treatment assignment is uncorrelated with the individuals' observable or unobservable characteristics, as everyone is equally likely to be selected, regardless of particular traits. Therefore, randomization creates a treatment group and a control group, from which a causal relationship between the intervention and impact indicator may be identified. The use of experiments in the Applied Social Sciences is not frequent due to issues ranging from ethical conflicts to financial and technical restrictions. In our hypothetical example concerning the ISE, drawing companies listed on the Brazilian Stock Exchange to participate in the ISE portfolio would configure an experiment along the lines discussed here. However, it is difficult to imagine this could happen in practice. The good news is that an experiment is not the only means for researchers to examine causal relationships.

Without an experimental design, the path to causal inference is to explore a **quasi-experiment** — or a **natural experiment** — which consists of a design in which treatment assignment is "as good as if it were random." There is a need to find an "exogenous variation" on the treatment variable, i.e., a source of variation that circumvents issues of self-selection of individuals to treatment to identify the causal effect of interest. Researchers can only identify this exogenous shock when they have in-depth knowledge about the institutional context linked to the phenomenon studied. Generally, the response lies in the eligibility criteria linked to the intervention. Other times, the source of exogeneity is not evident and is "between the lines" of the context. After identifying a natural experiment, an empirical strategy is required to perform causal inference, which we discuss below.

---

3   Angrist and Pischke (2009) simply and intuitively present the Rubin Causal Model, which addresses this problem by deriving the mean treatment effect of an intervention. The model uses the language of potential outcomes, where it is possible to decompose the average difference in outcomes observed between the untreated and treated groups into two parts: the treatment effect and the selection bias.

4   To avoid confusion regarding terminology, note that "random treatment assignment" differs from "random sampling." The first determines the individuals who will receive a given intervention, while the latter consists of randomly selecting a portion of elements belonging to a target audience. Therefore, obtaining a random sample from a population will not solve Endogeneity problems.

3. Identification Strategies

The identification strategy consists of the methodological approach that aims to overcome the Endogeneity problem and recover the causal interpretation of the relationship of interest. Quasi-experimental methods have received significant theoretical contributions over the last decades and are widely valued in empirical analyses due to the results' degree of reliability arising from methodological rigor. This section briefly presents the main methods used to identify causal effects to discuss the main hypotheses for identification and application contexts. Our objective here is not to develop each quasi-experimental method with statistical rigor but to approach the techniques from an introductory and intuitive perspective.

3.1 Difference-in Difference

Indeed, the identification strategy most frequently adopted in the mainstream literature is Difference-in-Difference (DiD). The application of the method is conditioned on the availability of longitudinal data for both treatment and control groups in at least two moments: before and after the treatment occurs. Specifically, it is necessary to observe the pre-intervention situation for both groups. The mean effect of interest is estimated by subtracting the differences between the dependent variable means of the treatment and control groups before and after the intervention. This double difference is responsible for assigning the name to the method. For the DiD estimator to identify the causal effect, one condition must be satisfied: that of **parallel trends**[5]. This identification hypothesis states that the impact indicator trajectories of the treatment and control groups would have similar dynamics in the absence of an intervention. Note that the condition **does not** require "the means of the impact indicator are equal between the groups" or that "the predetermined factors present equal means between the groups" before the intervention.

As it is impossible to observe the counterfactual situation of the treatment group (i.e., treated individuals not receiving treatment), it is easy to see that this identification hypothesis is not directly testable. Therefore, carrying out the sequence of differences previously explained does not ensure the identification of a causal relationship. For illustration purposes, the ISE example is applied again, considering that participating and non-participating companies are observed at two points in time, before and after the creation of the portfolio. Organizational performance could behave differently between treated companies even before joining the ISE, which indicates the presence of pre-existing differences in performance trajectories compared to companies outside the portfolio and, therefore, makes causal inference impossible.

In this sense, conducting empirical tests, such as verifying anticipatory behavior, which indirectly corroborates the hypothesis, is vital. Therefore, the DiD strategy is commonly explored in the context of observing multiple points in time. This is because if data are available at various times before and after treatment, we may verify whether there is evidence of pre-trends in the dependent variable. The DiD design that explores this expanded temporal dimension in a semi-parametric model in relation to the treatment variable is specifically known as an event-study[6]. In addition to tables, graphs are welcome and widely adopted to describe the indicator's trajectories and present the results.

---

5   There is also the hypothesis of the absence of anticipatory effects, commonly confused with parallel trends due to their close connection. For didactic reasons, we do not explicitly refer to this hypothesis. Roth et al. (2023) discuss both hypotheses in detail, the recent advancements in the DID literature, and future perspectives of the method.

6   Note that this design is different from the one classically applied in finance literature.

Another critical element is the possibility of differences in the timing of treatment adherence among treated individuals. Consequently, we can think of basically two distinct cases. The first is the canonical case, in which all treated units receive the intervention simultaneously. Complementary treatment is characterized by individuals adopting treatment that occurs progressively over time, known as staggered DiD. The second case deserves greater attention since it is reasonably common in designs in which DiD is adopted (Baker et al., 2022) and may be subject to bias when estimating the treatment effect. It occurs because heterogeneity at entry to the treatment — and potential heterogeneity of the treatment effect — might distort the weighting used in the estimator, introducing negative weights in the calculation of the mean treatment effect. As a consequence, the sign of the correlation may be erroneously inverted. A series of new DiD estimators correct the weights and generate reliable estimates of the causal effect to overcome such a difficulty. Figure 1 shows one of the results of Guimarães and Trevisan (2022), in which the authors use a regulation promulgated by the Brazilian Stock Exchange to estimate the effect of the mandatory separation of the CEO and chairman positions on a measure of shareholder value implementing three different estimators in a staggered event-study design. Figure 1 highlights the absence of anticipatory effects — statistically insignificant estimates and very close to zero before the separation — and indicates a positive and progressive effect on the impact indicator.
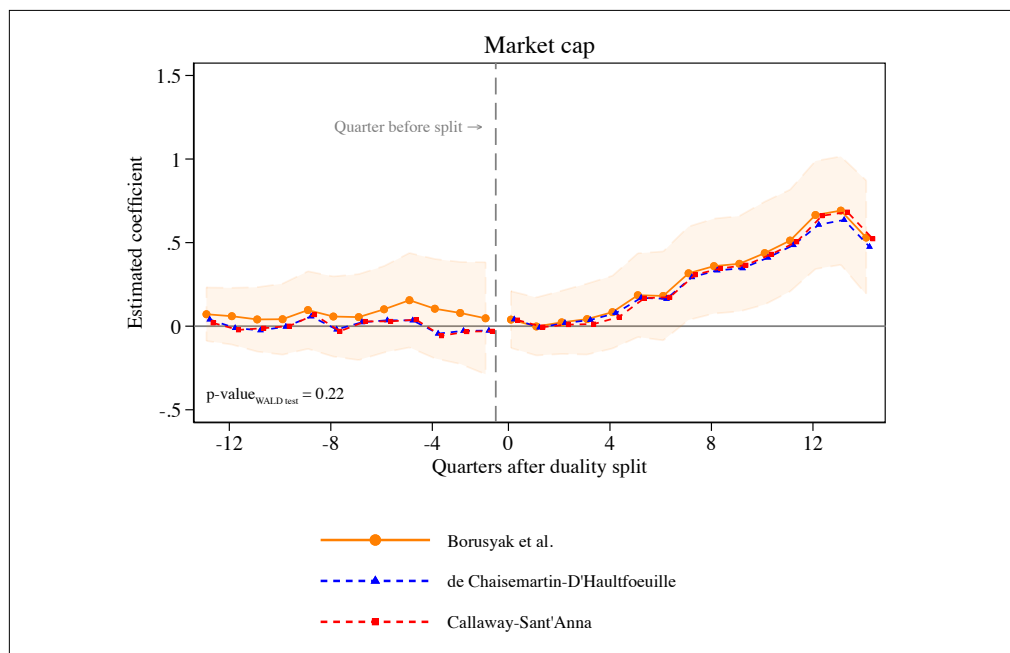

Figure 1. An example de staggered event-study design.

Source: Guimarães and Trevisan (2022).

It is worth highlighting that, to identify the causal effect in the DiD design, one must resort to motivations exogenous to individuals concerning their non-random treatment adoption. If treatment entry essentially involves self-selection of observational units, finding a suitable control group and extracting causality from the relationship of interest will be challenging. Therefore, it is interesting to explore rules, regulations, and guidelines of different natures, which constitute a component beyond the choice of individuals and generate variation over time and between groups. To better understand the application of the DiD method in Accounting research, we suggest reading Huang et al. (2020), Chircop et al. (2023), and Lin et al. (2019), which bring interesting applications to the topics of managerial litigation risk, tax avoidance, and impacts of IFRS adoption.

3.2 Instrumental Variables

Of all the quasi-experimental methods, the Instrumental Variable (IV) can be considered the most complex in identifying situations for its application. To present the method intuitively, we will recover the variables Y and D already defined in Section 2. We also consider the existence of another variable, W, which the researcher does not observe. The motivation is adopting a directed acyclic graph (DAG) used in every textbook presenting a canonical IV, as follows:
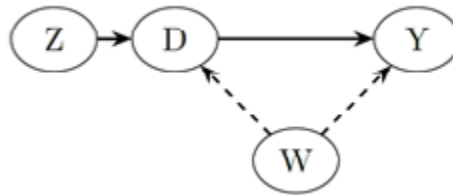

Figure 2. Gráfico direcionado acíclicon.

Fonte: elaborado pelos autores.

The causal relationship of interest is expressed by $D \rightarrow Y$. However, the W factor introduces a history of selection on unobservables since it confuses the relationship of interest by determining both the dependent and endogenous treatment variables ($D \leftarrow W \rightarrow Y$). Now, note the existence of factor Z and how it acts in the chain of causal effect expressed in the DAG: a variation in Z causes a variation in D, which causes a variation in Y. The most important perception here is that, even if Y varies when Z varies, the variation in Y **only occurs due to the variation in** D. In other words, the variable Z affects Y **exclusively through** D. This path is called **exclusion restriction**. The Z factor is an **instrumental variable** — or just an **instrument** —a term that gives its name to the identification strategy. Hence, a good instrument meets the exclusion restriction hypothesis (i.e., $Z \rightarrow D \rightarrow Y$) and is independent of other confounding factors. In this causal chain, you can think of factor D as assuming both the role of a mediating variable in the path $Z \rightarrow D \rightarrow Y$ and that of a collider in the direction $Z \rightarrow D \leftarrow W$.

However, why is it so difficult to find good instruments in practice? The reason is that they require a consistent theoretical and logical foundation to argue the validity of the exclusion restriction. This, by itself, is a good reason for researchers to avoid identifying a causal relationship by using IV, as such difficulty increases proportionally to the number of instruments. As in the case of parallel trends in DiD, the exclusion restriction hypothesis is not directly testable. Therefore, defending an instrument's validity requires applying several tests to verify the results' robustness and falsification of the empirical strategy. For example, it is essential to show that the instrument presents a strong correlation — in terms of statistical significance — with both the treatment variable (or **first stage**) and the dependent variable, in addition to presenting evidence of the absence of correlation with other relevant characteristics. However, one should remember that the first stage and the reduced form are insufficient to configure Z as a good instrument. Additionally, the researcher will need to tell an excellent story about it.

As highlighted by Cunningham (2021), the defense of a good instrument generally causes surprise to readers concerning the argumentation of its relationship with the dependent variable. Take as an example Bennedsen et al. (2007), who estimate the effect of nominating a family or an external CEO on corporate performance. Clearly, the decision to keep a family member in a company's management is endogenous — caused by previous performance problems, merger and acquisition decisions, and political connections, among several other unobservable factors. The authors propose using the "gender of the first child of the CEO who is about to be succeeded" as an exogenous variation to family succession. You may be wondering: How would the gender of the first child "influence" a company's performance? The elegance of the strategy is intrinsically linked to the context of this eccentric and peculiar reduced form. The authors explore unique data from family trees of corporate members and argue/show that the gender of the first child is a good instrument because it affects the probability of a family member being nominated **exclusively** due to birthright reasons (first-born males are more likely to "inherit" the position). Much of the previously mentioned paper is devoted to defending the validity of this exclusion restriction. On the other hand, it is improbable that this family trait is correlated with other determinants of company performance, as shown through the implementation of several robustness tests. This instrument's quasi-random attribute is what allows the causal relationship of interest to be identified.

IV applications are viable for both cross-sectional and panel data. Furthermore, it is common for research to use both "continuous" and "binary" treatments. Larcker and Rusticus (2010) carry out a literature review on the use of the method in research in the accounting field, discuss the non-triviality of its implementation, and promote a practical guide for its use. We warn that, whatever the situation in which a researcher believes in the possibility of using the IV strategy, there must be means (empirical and argumentative) of convincing the reader that the instrument meets the exclusion restriction since the use of unfit instruments can bias the estimator even more drastically than using a simple OLS. Some interesting examples of articles that adopt IV in the Accounting field are Fang et al. (2015) and Tseng (2022), which deal with the role of foreign investors in financial information reporting practices and the spillover effects of technologies on innovation.

3.3 Regression Discontinuity Design

Applications with the Regression Discontinuity Design (RDD) have been intensified in Applied Social Sciences due to the practicality of its implementation and convincing power to defend the results' internal validity. RDD explores a sudden change — of exogenous origin — in the probability of receiving a treatment, which occurs based on a specific value (cutoff point) of a quantitative variable (running variable). A quasi-experiment is created in the neighborhood/near this cutoff point so that the observational units within this neighborhood can be used as treatment and control groups due to their similarity in predetermined characteristics.

Thus, let us assume that tax authorities implement a tax inspection policy that institutes audits in companies that reach a specific value of annual net revenue, referring to the fiscal year preceding the year of the policy. In this case, the annual net revenue configures the quantitative variable (running variable), and the cutoff point would be the revenue limit that a company must reach to qualify for the inspection process. An RDD would enable comparing companies close to this threshold. Companies slightly above this threshold would be considered the treatment group, as they would be eligible for the tax inspection policy. In contrast, companies slightly below the threshold would be considered the control group, as they would not be selected for the audit.

Therefore, due to a **discontinuity** in the chance of receiving the treatment, the mean "local" treatment effect can be estimated, which is characterized as a **discontinuity** in the mean impact indicator precisely at the cutoff point. The identification hypothesis is that the mean of the dependent variable for the counterfactual group has a smooth transition at the cutoff. In other words, without the treatment, no discontinuity in the mean impact indicator should exist on the cutoff.

Not surprisingly, this hypothesis is not directly testable because it represents a situation that is impossible to observe. However, there are useful empirical tests to corroborate the support of this hypothesis and which are standard procedures in all studies using this identification strategy. The first is the balancing test, which checks the similarity between the treatment and control groups based on their observable characteristics. If the intervention is actually "as good as if it were random," the groups must be statistically equal regarding the means of the characteristics measured before the exogenous shock. Another critical test is to check if the running variable is manipulated. Intuition indicates that individuals could self-select to receive (or avoid) the treatment, manipulating the values of the running variable to (not) meet the eligibility criteria. To illustrate this point, consider the previous example of tax inspection policy. Suppose, alternatively, that tax authorities announce that companies that reach a minimum amount of net revenue at the end of the fiscal year will be subject to tax audits. In this case, companies will likely anticipate and adopt strategies to avoid exceeding the cutoff established by the policy not to be selected for the audit. If there is manipulation, the frequency of observations of the treatment and control groups at the cutoff will show a "jump," thus reflecting an endogenous behavior of the individuals and harming the causal inference.

The case in which all individuals receive treatment when they meet the eligibility criteria — the probability of receiving treatment jumps from 0% to 100% at the cutoff — is called a sharp design. This is the case in the previous example regarding the tax inspection policy. However, empirically, it is possible that not all individuals who meet the eligibility criteria actually receive the intervention. In other words, there may be situations in which there is endogeneity concerning the adoption of treatment. These cases configure the fuzzy design. To exemplify fuzzy RDD, let us continue using the previous hypothetical case. Suppose that, due to resource constraints, tax authorities cannot audit all eligible companies (i.e., those that have exceeded the minimum net income amount). This case configures a fuzzy design since the probability of being treated is less than 100% among eligible companies. In these cases, identifying the treatment effect on those treated uses the "opportunity" to receive the treatment as an IV to determine the probability of actually receiving the treatment. Note that the reduced form — the relationship between the impact indicator and the opportunity to receive treatment — represents the "treatment intention effect," an instrumental piece of information for those responsible for designing public policies or regulations. Estimating the parameter of interest can be done using non-parametric techniques, "optimal bands," and parametric techniques (OLS models). Cunningham (2021) appropriately discusses the current estimation methods and validation tests.

To illustrate the use of RDD in Accounting research, let us consider Lin et al. (2022), in which the impact of unionization on earnings management in the United States context is estimated. The study's motivation is related to the possibility that employees of unionized companies demand higher wages when companies have good financial results. Therefore, the previous study investigates whether unionization affects the practice of earnings management by reducing profits and, thus, mitigating the impact of higher compensation expenses. To achieve this objective, the authors explore a quasi-experiment originating in union elections, in which unionization is instituted if the company obtains at least 50% plus "1" of the votes. Hence, the unionization event is almost random for companies in which the votes were close. The authors compared companies that were successful in unionization by a margin close to 50% with those that were also unsuccessful by a margin close to this limit (sharp drawing). Panel A of Figure 3 graphically presents one of the article's results. Note that the curve that represents the mean behavior of earnings management presents a discontinuity precisely at the cutoff, indicating that, after unionization, there is a reduction in discretionary accruals. This result signals that companies use managerial discretion to report lower profits, trying to shield themselves from workers' demands for higher wages. Panel B performs a placebo test, simulating that the actual cutoff would be at the "0.45" point instead of "0.50". The graph suggests that there is no discontinuity at this point, thus validating the strategy of identifying the causal effect of unionization.
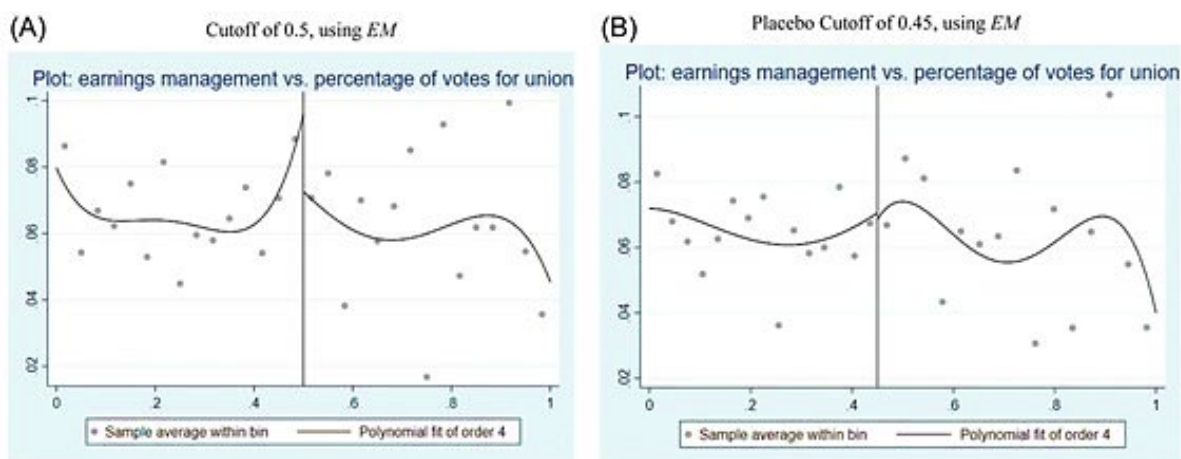

Figure 3. Example of sharp RDD.

Source: adapted from Lin et al. (2022).

In short, for RDD to be applicable, there must be an exogenous and sudden variation in the chance of receiving treatment based on a specific value of a quantitative index. A significant advantage of the method is that it does not exclusively depend on longitudinal data to be implemented. We recommend reading Joshi (2020) and Fan et al. (2021), in which RDD is used to assess the effects of regulation related to the reporting of tax information on companies' tax avoidance and the impacts of proposals related to corporate governance on earnings management.

4. Final Considerations

One of the most admirable aspects of scientific thinking is its premise of allowing it to break paradigms and continually renew itself. As researchers, we need to be receptive to "the new" to continue learning and contributing to advancing knowledge about phenomena in which we are involved. In this editorial, we highlight an epistemological issue in the context of quantitative empirical research in the accounting field, seeking to infer a causal relationship. In this sense, we would like to encourage some reflections: Do we adopt the appropriate methodological tools to seek answers to our research questions? Are we adequately updated, and have we learned about the methods available? From our point of view, such knowledge is not yet fully disseminated and consolidated, and there is still a long way to go.

Quantitative methods for causal inference, especially at a national level, need to be rethought: a structural break in the *modus operandi* of Accounting research is needed. A change in standards regarding the mainstream (quantitative) accounting literature has been seen among the leading journals concerning the demand for studies implementing more rigorous methods to identify causal relationships and explore quasi-experiments and controlled experimental designs. To converge and interact with the mainstream, we need to appropriate such methods. Therefore, this editorial is expected to clarify some introductory questions regarding quasi-experimental econometric techniques and instigate the academic community to "rediscover" it and learn about its potential use.

References

Ahern, K. R. & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. The Quarterly Journal of Economics, 127(1), 137–197. https://doi.org/10.1093/qje/qjr049.

Angrist, J. D. & Pischke, J. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? Journal of Financial Economics, 144(2), 370–395. https://doi.org/10.1016/j.jfineco.2022.01.004.

Ball, R. & Brown, P. (1968). An empirical evaluation of accounting income numbers. Journal of Accounting Research, 6(2), 159–178. https://doi.org/10.2307/2490232.

Beaver, W. (1968). Market prices, financial ratios, and the prediction of failure. Journal of Accounting Research, 6(2), 179–192. https://doi.org/10.2307/2490233.

Bennedsen, M., Nielsen, K. M., Perez-Gonzalez, F., & Wolfenzon, D. (2007). Inside the family firm: The role of families in succession decisions and performance. The Quarterly Journal of Economics, 122(2), 647–691. https://doi.org/10.1162/qjec.122.2.647.

Chircop, J., Fabrizi, M., Malaspina, P., & Parbonetti, A. (2023). Anti-mafia police actions, criminal firms, and peer firm tax avoidance. Journal of Accounting Research, 61(1), 243–277. https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-679X.12455.

Cunningham, S. (2021). Causal inference: The mixtape. Yale university press.

Fan, Z., Radhakrishnan, S., & Zhang, Y. (2021). Corporate governance and earnings management: Evidence from shareholder proposals. Contemporary Accounting Research, 38(2), 1434–1464. https://onlinelibrary.wiley.com/doi/abs/10.1111/1911-3846.12640.

Fang, V. W., Maffett, M., & Zhang, B. (2015). Foreign institutional ownership and the global convergence of financial reporting practices. Journal of Accounting Research, 53(3), 593–631. https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-679X.12076.

Gow, I. D., Larcker, D. F., & Reiss, P. C. (2016). Causal inference in accounting research. Journal of Accounting Research, 54(2), 477–523. https://doi.org/10.1111/1475-679X.12116.

Guimarães, R. & Trevisan, G. (2022). Mandatory CEO Non-Duality, Managerial Agency, and Shareholder Value. Working Paper, SSRN. https://ssrn.com/abstract=4122646.

Huang, Y., Li, N., Yu, Y., & Zhou, X. (2020). The effect of managerial litigation risk on earnings warnings: Evidence from a natural experiment. Journal of Accounting Research, 58(5), 1161–1202. https://doi.org/10.1111/1475-679X.12336.

Joshi, P. (2020). Does private country-by-country reporting deter tax avoidance and income shifting? evidence from beps action item 13. Journal of Accounting Research, 58(2), 333–381. https://doi.org/https://doi.org/10.1111/1475-679X.12304.

Larcker, D. F., Richardson, S. A., & Tuna, I. (2007). Corporate governance, accounting outcomes, and organizational performance. The Accounting Review, 82(4), 963–1008. https://doi.org/10.2308/accr.2007.82.4.963.

Larcker, D. F. & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. Journal of Accounting and Economics, 49(3), 186–205. https://www.sciencedirect.com/science/article/pii/S0165410109000718.

Lin, G., Brown, A. B., & Wu, W. (2022). The effect of unionization on firms' earnings management and employee compensation: Evidence from a regression discontinuity design. Journal of Business Finance Accounting, (pp. 1–28.). https://doi.org/10.1111/jbfa.12665.

Lin, S., Riccardi, W. N., Wang, C., Hopkins, P. E., & Kabureck, G. (2019). Relative effects of ifrs adoption and ifrs convergence on financial statement comparability. Contemporary Accounting Research, 36(2), 588–628. https://onlinelibrary.wiley.com/doi/abs/10.1111/1911-3846.12475.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-indifferences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008.

Tseng, K. (2022). Learning from the joneses: Technology spillover, innovation externality, and stock returns. Journal of Accounting and Economics, 73(2), 101478. https://www.sciencedirect.com/science/article/pii/S0165410122000015.

Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage learning.