



From Meaningful Data Science to Impactful Decisions: The Importance of Being Causally Prescriptive

RESEARCH PAPER

VICTOR S. Y. LO 

DESSLAVA A. PACHAMANOVA 

*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

This article proposes a framework for transition from traditional data science, where the focus is on extracting value from available data, to goal-driven analytical decision-making, where the business objective is defined first, through integration of various analytical techniques in a common setting. We discuss the link between predictive analytics and prescriptive analytics in the context of formulating the problem and assert that all prescriptive analytics problem formulations assume a causal link between decisions and outcomes. We emphasize the role of predictive analytics and causal inference in specifying the causal link between decisions and outcomes accurately and ultimately in aligning the analysis with the business objectives. We offer practical examples that integrate various required analytics tasks and describe scenarios where causal inference is required versus not required.

CORRESPONDING AUTHOR:

Victor S. Y. Lo

Senior Vice President of
Data Science and Artificial
Intelligence, Workplace
Investing, Fidelity Investments,
Boston, MA 02210, US

victor.sy.lo@gmail.com

KEYWORDS:

prescriptive analytics;
predictive analytics; causal
inference; data science;
decision-making

TO CITE THIS ARTICLE:

Lo, VSY and Pachamanova, DA. 2023. From Meaningful Data Science to Impactful Decisions: The Importance of Being Causally Prescriptive. *Data Science Journal*, 22: 8, pp. 1–18. DOI: <https://doi.org/10.5334/dsj-2023-008>

Rapid growth in data, industry demand, and technology have created a golden age for data science. More than ever, there is widespread recognition that data and analytics bring business value by enabling informed decisions. The analytics underlying business decisions is often categorized into three types: descriptive, predictive, and prescriptive (Figure 1). The foundational level—descriptive—involves data summaries, visualizations, and observation. The middle level—predictive—involves building models to explain and predict future behavior, with significant contributions from the statistical, machine learning, and econometrics fields. The top level—prescriptive—operationalizes the insights from the previous two levels, taking into consideration business constraints and often utilizing advanced prescriptive analytical methods from fields like operations research and system dynamics; see also Rose (2016), LaRiviere et al. (2016), Frazzetto et al. (2019), Poornima & Pushpalatha (2020), Bertsimas & Kallus (2020), Delen (2020), and Lo (2020).

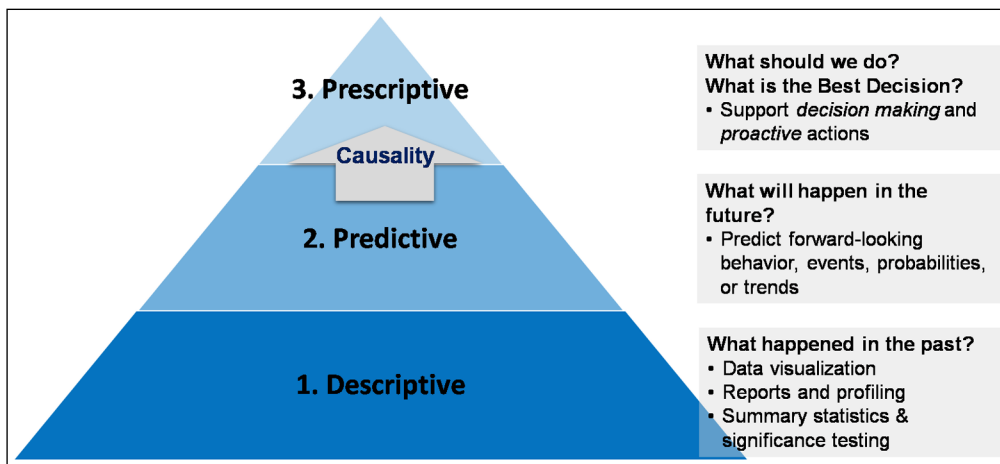


Figure 1 Three types of analytics (Lo 2020).

Successfully integrating all three levels of analytics in the context of business process improvement requires combining data science with decision science (Bobriakov 2019). As argued in de Langhe and Puntoni (2021) and Bojinov, Chen, and Liu (2020), data-driven strategies that are not decision-driven strategies can be seriously flawed. From a methodological point of view, transitioning from data science to decision science requires (1) a holistic view of the process—from data to descriptive, predictive, and prescriptive analytics—and (2) generating the correct inputs through predictive analytics to inform prescriptive analytics. This article analyzes this transition through the lens of *causality*.

The study of causality, or cause-and-effect relationships, has been around for centuries in multiple fields. While correlation or association between two variables can be observed from available data, it is only what we “see” about how they are related and is not about the effect of what we “do” (or manipulate by setting a specific value) to one variable on the other variable Pearl & Mackenzie (2018) (Illari & Russo 2014). For example, if the displayed times on the watches of the two authors always differ by a minute, the two times are clearly correlated and are both driven by the official time (a confounder). If we move up the time on one watch, there is no reason to expect the other watch will automatically be impacted, as they are not causally related. Now consider a different example: a student’s exam score and its relationship to the level of preparation of the student for the exam. It may be reasonable to assume that better preparation for the exam will result in a better exam score on average, that is, the relationship between the level of student preparation and exam score would be causal. If this is the case, an intervention or manipulation (preparation for the exam) can *causally* change the probability distribution of a system response or outcome (exam score). By maximizing preparation, the student can maximize the exam score.

Causal inference is a set of methodologies based on experimental or observational data to learn about cause-and-effect relationships by eliminating or reducing confounding effects, allowing us to assess what decision or action can cause a desirable outcome and quantify the causal relationship. Understanding the cause-and-effect relationship enables a better decision to be made. As we will point out in this article, however, in practice, one needs to carefully identify the ultimate goal and the levers to accomplish this goal.

In the student exam preparation example, “preparation” is a vague term. What are the actual levers a student has to prepare for an exam? (These levers are represented by specific terms in prescriptive analytics.) Does *preparation* mean “time to prepare”? Can we reasonably expect that the time a student spends preparing will increase the final exam score? Even if we assume it will, defining the goal is not obvious. Is the student’s goal to perform well on the exam? Or is there an overarching goal, like performing well in the course? If the ultimate goal is to perform well in the course, the student’s time may need to be split between different course deliverables, of which the exam will be one but not necessarily the only one. Maximizing the time spent studying for the exam may not necessarily result in accomplishing the goal of performing well in the course overall, even if the causal relationship between studying for the exam and achieving a desirable outcome (a high exam score) is understood and can be estimated. Extending this simplified example to business situations, in this article, we will argue that specifying the ultimate goal carefully and mapping out the correct causally prescriptive methodologies in the process for achieving that goal is paramount to impactful decision-making.

The decision theory literature from psychology and philosophy considers causal knowledge essential to decision-making, as humans choose from available options that *cause* the outcomes that they desire; see Hagmayer & Sloman (2005), Sloman & Hagmayer (2006), Joyce (2008), and Hagmayer & Fernback (2017). While these scholars are mostly concerned about how decisions are made by humans, we are interested in how decisions *should* be made in a prescriptive fashion; see Sloman (2005: ch.7) for comments on the difference. Surprisingly, many paradigms used in prescriptive analytics from decision science fields, such as operations research, operations management, and industrial engineering, do not include explicit considerations for causality or causal inference. Considerations for causality are incorporated in paradigms from other analytical fields. For example, Attri, Dev, and Sharma (2013); Ali, Sorooshian, and Kie (2016); Chauhan, Singh, and Jharkharia (2018); and Sorooshian, Tavana, and Ribeiro-Navarrete (2023) describe the DEMATEL and ISM methodologies, which address causal relationships between variables for decision-making based on expert judgment. However, such methodologies do not optimize specific outcomes as in traditional prescriptive analytics formulations and do not explicitly incorporate estimation procedures for causal relationships based on scientific experiments (such as A/B testing and clinical trials) or observational data.

Our contribution is to introduce a practical seven-step causal prescriptive analytics framework that blends important concepts and methodologies from several different analytical fields and specifically addresses the question of when causal inference is necessary to support prescriptive analytics paradigms. We illustrate the usage of the framework with examples.

2. INTRODUCING THE CAUSAL PRESCRIPTIVE ANALYTICS FRAMEWORK

A common characteristic in many decision problems is that they can be represented as optimization models, making optimization modeling an important prescriptive analytics tool. There are three parts to the classical optimization paradigm: (1) decision variables (actions or quantities under the control of the decision maker), (2) objective function (goal expressed in terms of the decision variables), and (3) constraints (requirements on the decision variables). Optimization formulations play an important role in prescriptive analytics, allowing for situations to be described in a common language that can then be passed to optimization solvers to obtain an optimal policy.

Because the predictive analytics and optimization communities have traditionally been separated, we find that there are important steps missing from applying prescriptive methodologies like optimization in practice. The practical causal prescriptive analytics framework we propose below replaces the standard optimization paradigm with **seven key questions** that need to be answered to understand the *end-to-end process* needed to support informed decision-making (Figure 2):

1. What is the **objective or goal**? That is, what are you trying to achieve?

The first question for any decision problem is to know what the ultimate goal (Z) is. Although this might seem obvious, in our experience, many analytics and data science projects do not start with a goal or the goal is poorly defined. For impactful decision-making, we propose defining a goal that one wants to achieve rather than finding purpose for available

data or focusing on exciting methodologies; see Keeney (1996) for various considerations in objective setting.

2. What are the **outcomes** that help achieve your objective?

Defining the outcomes that can be achieved and how they relate to the ultimate goal is critical for understanding whether the goal (Z) is achievable. These immediate outcomes (Y) are driven by actions or decisions (X in Question 3) and are expressed as a function of these decisions.

3. What are the **decision variables** (options, actions, treatments, or interventions)?

It is important to know what kind of decision (X) one could possibly make in order to influence the outcome (Y from Question 2) so as to ultimately achieve one's goal (Z from Question 1). The decision variables should be defined so that they are actions *under the decision maker's control* and so that their *causal* connection to the outcomes (Y) that help achieve the goal (Z) is well understood.

4. What is the **available information**, such as data, insights, and models?

Existing information (I) may exist that can help influence the decision (X) or can be used to define the relationship between X and Y. This information can be based on domain knowledge about the relationship between X and Y or availability of experimental or observational data from past decisions and outcomes that can be used to infer the causal relationship between X and Y. Taking advantage of such information allows one to express the causal relationship between X and Y correctly (see Question 5).

5. What is the relationship between X and Y?

Understanding the nature of the relationship between X and Y allows one to build models to represent it accurately. This representation is often estimated through predictive analytics or causal inference techniques, unless the relationship is already known. It is what enters the prescriptive model and is used to identify desired outcomes. We will further address the significance of this step in section 5.

As illustrated in Figure 2, in practice, Questions 1–5 are resolved in an iterative manner. One may need to revisit the specification of goals, outcomes, and actions multiple times to define them in a way that allows for taking advantage of available information and building an accurate representation of the relationship between X and Y.

6. What **constraints** need to be taken into consideration?

There are often constraints that limit the range of potential decisions X. Some of these constraints are physical (e.g., one cannot market to a negative number of customers), while others are imposed by business practices.

7. What is an **appropriate solution** that achieves the goal?

After formulating the problem in the steps above, the final step is to determine an appropriate solution. There may be multiple combinations of methods for solving the prescriptive analytics problem. The choice of method combination and solution depends on available data, time, and resources.

The seven questions above extend and enrich the three components of the classical optimization paradigm to emphasize the implementation and usability of model insights for impact. As depicted in Figure 2, the framework is a cycle: it starts and ends with the stated goal, making sure that the whole process is aligned with accomplishing the goal. Note that the choice of decision variables (X) drives the immediate outcome (Y), which leads to the ultimate goal (Z). Decision-making in this framework is therefore assumed to be **causal** by nature. However, whether X and Y are defined accurately to support Z (Questions 2 and 3) and whether the correct relationship between X and Y is estimated (Question 5) can have a significant effect on the final goal. In other words, before employing exciting prescriptive analytics methodologies, one needs to understand whether making a decision (X) would affect the goal (Z) in a predictable *causal* manner through the function that is used to describe their relationship and whether this will ultimately support Z. In some cases, causal inference is needed to estimate the relationship between X and Y, while in others, it is not. Understanding the situations in which causal inference is necessary is critical for selecting business-relevant optimal solutions (Question 7).

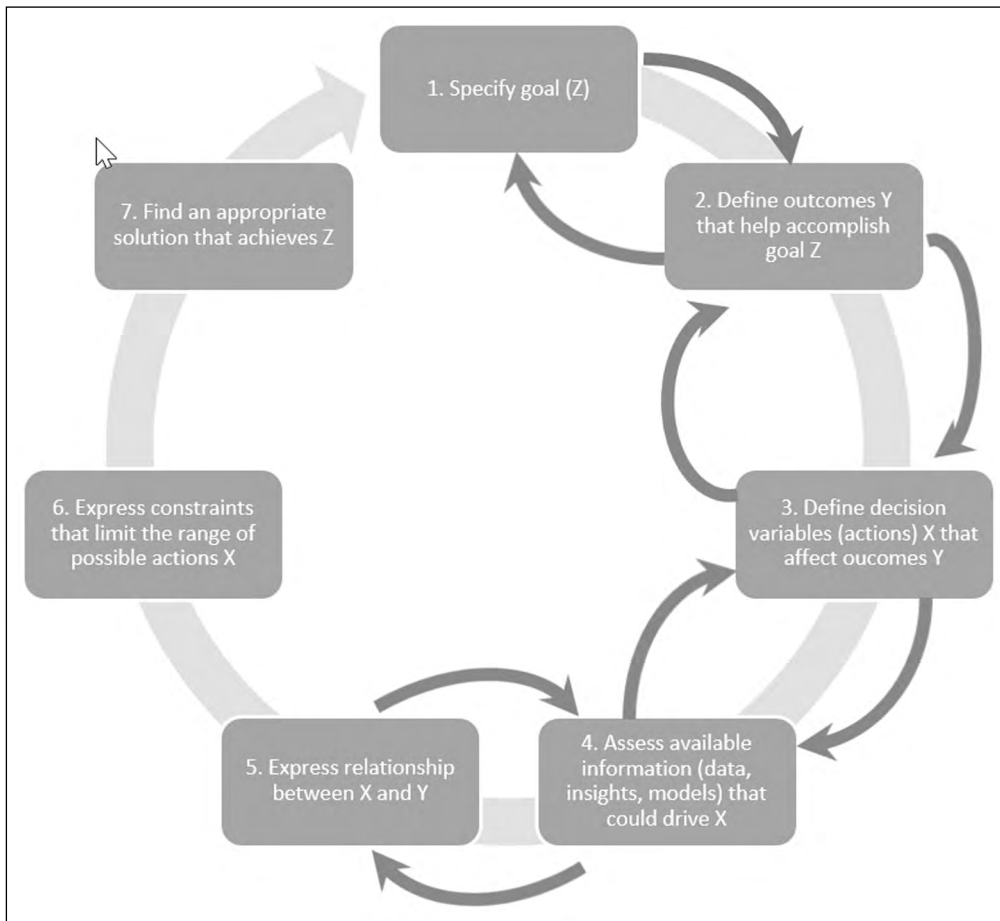


Figure 2 Proposed causal prescriptive analytics framework.

Let us illustrate this point and the causal prescriptive analytics framework with a common application: direct marketing. Direct marketing involves intervention or treatment, such as an email, a direct mail, a web advertisement, or a phone call to a customer, that aims to maximize a call to action, such as a product purchase. Marketers typically deal with multiple treatments and/or multiple products and have a fixed budget. The goal is to match each customer with the right treatment and/or product so as to maximize a business outcome, such as sales or expected profit.¹ The seven-step process for this example is as follows (Figure 2):

1. What is your **objective or goal**? That is, what are you trying to achieve?

From the perspective of the business, the overarching goal is to maximize profit (Z), which means maximizing the customer purchase response *that results from direct marketing*. This nuance is important, as ideally only customers who are likely to change their behavior and purchase due to direct marketing should be treated (targeted) in order for marketing spend to be allocated purposefully; see, for example, Lo (2002; 2008) and Lo & Pachamanova (2015) for additional discussion. Well-run businesses can incent pursuing desirable goals by aligning performance metrics with these goals. In this context, accounting for the additional number of customers who purchase (or additional expected revenue) with treatment (e.g., email) relative to no treatment at all or a business-as-usual intervention that has been used in the past are examples of metrics that create such incentives. These metrics directly relate to the profitability of the direct marketing campaign and the business's bottom line (Z).

2. What are **outcomes** that help achieve your objective?

Profitability is a function of incremental sales volume, which is directly linked to the probability that an individual customer changes behavior as a result of the direct marketing campaign. The outcome (Y in Figure 2) that directly affects the goal (Z) is therefore the total incremental difference in individual customer purchase probabilities (also known as *lift*) as a result of the marketing campaign.

¹ In addition to being a key tool for many for-profit organizations, direct marketing can also be generalized to other similar situations, such as contacting the right donors for donations in a fundraising program for a nonprofit organization.

3. What are the **decision variables** (i.e., options, actions, treatments, or interventions)?

The actions or decision variables are whether or not to send a particular offer (treatment) to each customer. If there are multiple treatments (or products), one can define the decision variables (X) to be binary (0 or 1) to correspond to the treatment being considered for each customer. If there are M treatments and N customers, there are a total of MN binary decision variables.

4. What is the **available information** (e.g., data, insights, or models) to influence your decision?

Companies often have data on whether or not customers have purchased a particular product in the past. They also have data on whether or not a customer has been targeted with a particular campaign. However, the data set needs to be assessed carefully in the context of understanding whether an action can lead to customer response. In this context, one needs prior marketing campaign data based on a *randomized controlled trial* (also known as A/B testing in business applications) that contains response data (whether the customer purchased or not) to each treatment as well as demographic data as covariates for model estimation. Such data can be used to test scientifically and measure the effectiveness of a treatment, allowing for evaluation of the incremental probability that the customer will purchase as a result of the marketing campaign.

5. What is the relationship between X and Y ?

With the availability of the right data set in Step 4, one can estimate the incremental probability of response for a particular customer using causal inference techniques; see, for example, Kane, Lo & Zheng (2014). The total change in response probability because of the marketing campaign can be estimated as a sum-product of the individual response probabilities (lift) with the binary decision variables of whether or not a particular treatment is applied to a particular customer; see Lo & Pachamanova (2015) and Appendix A for a mathematical formulation and an empirical example.

6. What **constraints** need to be taken into consideration?

The main business constraint is that the total cost of all treatments cannot exceed a fixed marketing budget. There could be additional business constraints, such as the limitation that each customer should only receive at most one treatment in this marketing campaign. Physical constraints include restrictions on the decision variables; they need to take only values 0 or 1 for the formulation of the problem to be meaningful.

7. What is an **appropriate solution** that achieves the goal?

After going through Steps 1–6, one needs to take a holistic look at the path to find a solution and evaluate whether it ultimately addresses the goal in Step 1. In this example, since the goal is to maximize profitability due to direct marketing, a binary integer optimization program can be set up to maximize profitability by determining the right treatment for each customer (i.e., determining the values for the decision variables X that lead to the highest outcome Y , which was determined to lead to the desired goal Z). As a key input to the optimization model, uplift modeling or conditional average treatment effect (CATE) techniques can be employed to estimate the treatment effect as a function of available covariates for each treatment in Step 5, enabling the prediction of the effect of each treatment over control (e.g., no action) at the individual or subgroup level. The combination of uplift modeling and constrained optimization requires an integrated predictive and prescriptive analytics approach; see also Lo (2002; 2008); Kane, Lo & Zheng (2014); Lo & Pachamanova (2015); Pachamanova, Lo & Gulpinar (2020); and Appendix A.

The causal prescriptive analytics framework can be illustrated through a directed acyclic graph (DAG) (Figure 3). The box around the decision X represents the possible constraints that are “boxing” the feasible values of X .

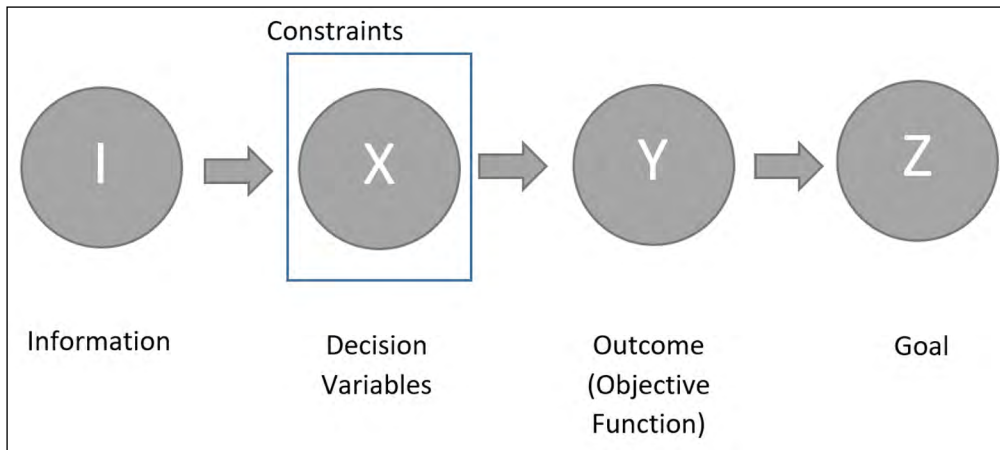


Figure 3 Directed acyclic graph (DAG) representation of the proposed causal prescriptive analytics framework.

3. APPLICATIONS OF THE CAUSAL PRESCRIPTIVE ANALYTICS FRAMEWORK

Many important problems can be addressed in practice with the causal prescriptive analytics framework introduced in section 2. We list several examples below.

Vehicle routing. Vehicle routing is a common and very difficult problem in prescriptive analytics, with applications ranging from delivery service routing to bus routing (Routing Challenge 2021; Bertsimas et al. 2020; Davenport 2013). For instance, routing a fleet of school buses requires that each student be picked up and delivered to the school within particular time windows, while minimizing total travel time or distance, and is subject to various constraints, such as bus capacity. It can be formulated as a constrained optimization problem by representing all location points where students wait as vertices in a graph and assigning binary decision variables X to correspond to the actions of whether or not to use particular arcs in the graph (Feillet 2010). The outcome Y (e.g., total travel distance) can be expressed as a sum-product of arc distances and decision variables representing whether certain arcs are used. Constraints include the capacity of buses, the fact that pickup and drop-off need to happen within particular time windows, and the fact that the arcs selected in a bus route need to form a continuous path with predetermined start and end points.

Workforce scheduling. Workforce scheduling problems appear in multiple business contexts, from large retailers to call centers to hospitals. The main goal is to set a timetable assigning a particular number of employees to shifts (decisions X) so that employee preferences or labor law constraints are taken into consideration while meeting business demands and minimizing total cost ($Y = Z$); see Daskin (2010: ch.7–8) and Koole (2013: ch.5–6).

Inventory management. A factory responsible for producing a product needs to order raw materials with the right amount (decision X) in order to minimize the total cost (Z), which is the sum of ordering cost ($Y1$) and inventory holding cost ($Y2$).

Portfolio construction. The portfolio optimization problem has been a central problem in quantitative investments since Markowitz (1952) proposed considering the trade-off between expected reward and risk to determine the optimal portfolio allocation. Given estimates of future asset expected returns and risk, the goal (Z) is to maximize the future value of the funds, which can be translated as finding the asset weights (X) that maximize the expected portfolio return (Y) for a given target level of portfolio risk.

Pricing. The prices of products and services are often determined by several departments in a company, such as accounting, marketing, and product managers. There is a minimum price that can be charged based on the cost of the product, but to set the final price, one needs to estimate the supply and demand for the product at any given price. If the price is set too high, the demand would be lower, but the profit margin per item would be higher, and vice versa. The goal is to determine a price level (X) that results in the maximum profit; see Lo (2020).

Customer retention. This classic problem involves identifying the best customers for special customer retention efforts in order to improve profitability by balancing the cost and benefits of retention. This problem can be formulated as follows: the objective is to maximize profitability (Z) through customer retention (Y) by attempting to retain the appropriate customers (X).

Employee acquisition. The optimal number of employees to be hired is a common but not simple problem with the goal of maximizing overall profitability. For example, in a department store, the employee acquisition decision involves how many additional sales employees to recruit (X) in order to maximize overall incremental profit (Z), defined as sales revenue (Y) minus employment cost.

Digital health. Wearable devices not only report patient vitals but can also be used to provide health-related recommendations to patients. These devices are sometimes freely offered by employers or insurers. Relevant decisions include which of a set of messages to display (and when) for each individual in order to achieve positive outcomes, such as minimizing emergency room visits and medical costs (Menictas et al. 2019; Carpenter et al. 2020).

Personalized medicine. The National Institutes of Health (NIH) and Food and Drug Administration (FDA) jointly proposed personalized medicine in Hamburg and Collins (2010), with the goal of providing patient-level individualized treatment as opposed to the typical one-size-fits-all medical treatment. Within the causal prescriptive analytics framework paradigm, this problem can be formulated with decision variables (X) that correspond to individuals to be selected to receive treatment so as to maximize effectiveness (Y) and ultimately improve population health (Z). This emerging field has many challenges to overcome, including measurement and optimization.

Government policies. Economic and health care policies have a wide impact on the community with objectives such as improving health, cost control, and improving the economy. Examples include determining the necessary level of interest rate (X) to optimize consumer and business response (Y) and ultimately improve overall economic measures (Z) or tuning health care policy parameters (X) to treat hospitals serving vulnerable populations equitably (Y) and ultimately reducing disparity in treatment in the population (Z). Table 1 contains a technical summary of these examples. (The “Coefficient, C” column and the classification of examples by causal inference requirement into Panels A and B of Table 1 will be explained in section 5.)

Table 1 Examples of problems where the causal prescriptive analytics framework can be applied.

PANEL A: CAUSAL INFERENCE NOT REQUIRED				
PROBLEM	DECISION, X	COEFFICIENT, C	IMMEDIATE OUTCOME, Y	ULTIMATE OBJECTIVE, Z
Vehicle routing	Selection of arcs	Arc distance	Total travel distance	Travel distance or cost, to be minimized
Workforce scheduling	Assignment of employees to shifts	Staffing cost per employee	Total cost	= Y, to be minimized
Inventory management	Quantity of raw materials ordered at each time	Holding cost per unit and cost per order	Ordering cost and holding cost	Total cost, to be minimized
Portfolio construction	% allocation to each stock	Individual stock returns	Monthly portfolio return	Long-term return, to be maximized
PANEL B: CAUSAL INFERENCE REQUIRED				
Direct marketing (see Appendix A for details)	Assignment of treatment to each customer	Lift in purchase probability due to direct marketing	Incremental sales due to direct marketing	Incremental profit due to direct marketing, to be maximized
Pricing	What price to set	Sales volume	Sales revenue	Profit = sales revenue – variable cost, to be maximized
Customer retention	Attempt to retain which customer	Change in retention rate due to retention program	Retained or not	Profit = predicted revenue from future sales *P(retention) – cost of retention, to be maximized
Employee acquisition	Number of sales agents to recruit	Total sales volume	Total sales revenue	Profit = estimated sales revenue (Y) – cost of total employment, to be maximized
Digital health	Message to show to each individual	Message-specific health outcome	Individual health outcome	Employer-level health cost, to be minimized
Personalized medicine	Who to receive treatment	Treatment effectiveness	Individual health outcome	Population health, to be maximized
Health care policy	Introduce the policy or not	Population readmission rate	Population readmission rate	Health care cost, to be minimized
Economic policy	Interest rate level	Consumer and business responses	Consumer and business responses	Overall economic measure, to be improved

The different stages of the causal prescriptive analytics framework require methodologies from multiple analytical fields. Most generally, they can be separated into three categories: predictive analytics, causal inference, and optimization.

Predictive analytics. Information (I) in Figure 3 may be provided directly or may need to be estimated based on historical data using statistical analysis, predictive analytics, or machine learning models. The range of methods includes point and interval estimation, regression-based analysis, econometric time series analysis, decision tree, random forest, gradient boosted tree, Bayesian analysis, and deep learning; see, for example, Mills & Markellos (2008); Freedman (2009); Hastie, Tibshirani & Friedman (2009); Theodoridis (2015); Gelman et al. (2013); and Gelman, Hill & Vehtari (2021).

Causal inference. While predictive analytics is about predicting an outcome (Y) as accurately as possible using available features (X) where the metric of interest is typically the conditional expected value $E(Y|X = x)$ or conditional probability $P(Y = 1|X = x)$, causal inference is for measuring the effect of a cause on the outcome. The former can be directly addressed by modeling $E(Y|X = x)$ using statistical or machine learning methods for supervised learning to approximate the functional (not necessarily causal) relationship between Y and X. The latter would require specific techniques from the field of causal inference; using the do-calculus notation from Pearl (2000) and Pearl, Glymour, and Jewell (2016), the causal effect of X on Y is denoted by $E(Y|do(X = a)) - E(Y|do(X = b))$, where the do-operator is to indicate that an intervention is acted on the value of X as opposed to simply observing the value of X in a conditional expectation. To measure the causal impact of a decision or intervention, randomized controlled trial (RCT), where treated and untreated units are randomly split, is often regarded as the gold standard and is recommended to be used whenever possible; see Salsburg (2001: ch.5), Freedman (2009: ch.1), Glennerster & Takavarasha (2013), Imbens & Rubin (2015), Pearl & MacKenzie (2018), Leigh (2018), Rosenbaum (2019), and Thomke (2020). In situations where RCT is unavailable because experimentation is too difficult or too costly to achieve, confounding can be reduced by causal inference techniques for observational data. This itself is a large field that cuts across several academic disciplines and is out of scope for this paper, and thus we provide some key citations below for each major school of thought:

1. The potential outcomes or counterfactual approach and the related propensity score matching from the field of statistics, where it is assumed that hypothetical outcomes exist under the treated and untreated scenarios for each analysis unit and a single propensity score, defined as the probability that each analysis unit belongs to treatment as a function of confounders, can be applied to match treatment and control groups as much as possible; see Rubin (2006); Rosenbaum (2002; 2010; 2019); Imbens & Rubin (2015); Hernan & Robins (2020); and Dominici, Bargagli-Stoffi & Mealli (2021) for a description of multiple variations of this technique widely applied to social and medical sciences.
2. The probabilistic graphical method from artificial intelligence based on the Bayesian network is a unique approach, as it is designed to construct and estimate the causal relationships (or DAG) through structure learning, even without detailed domain knowledge of which variables may be causing which, thus allowing us to answer causes of effects in addition to effects of causes; see Pearl (2000; 2012); Spirtes, Glymour & Scheines (2000); Koller & Friedman (2009); Scutari & Denis (2015); Pearl, Glymour & Jewell (2016); and Peters et al. (2017).
3. The natural and quasi-experimental methods from economics and social sciences employ a collection of econometric methods, including difference-in-difference, panel data analysis, regression discontinuity design, and instrumental variable; see Angrist & Pischke (2009), Morgan & Winship (2015), Dunning (2016), and Reichardt (2019).
4. Tracing the cause-and-effect path through causal mechanisms from epidemiology and social sciences and identifying direct and indirect effects of treatments; see Iacobucci (2008), Hayes (2013), and Vanderweele (2015).

Optimization. Selecting the best decision to maximize or minimize the ultimate objective subject to constraints is a constrained optimization problem that falls in the realm of operations research and industrial engineering. Techniques available for solving constrained optimization include classic mathematical programming techniques such as linear programming, integer programming, and nonlinear programming (see Bertsimas & Tsitsiklis (1997), Papadimitriou & Steiglitz (1998), and Williams (2003)); optimization under uncertainty methods such as stochastic programming and robust optimization (see Wallace & Ziemba (2005), Cornuejols & Tutuncu (2007), Fabozzi et al. (2007), and Ben-Tal et al. (2009)); and multiobjective optimization methods (see Antunes, Alves & Climaco (2016) and Kaliszewski, Miroforidis & Podkopaev (2016)). Reinforcement learning methods integrate stochastic and dynamic optimization over multiple stages with statistical estimation; see Powell (2011), Sugiyama (2015), Kochenderfer (2015), and Sutton & Barto (2018).

This overview of techniques indicates that the appropriate solution to an optimal decision-making problem is often *highly multidisciplinary*. Therefore, we note that Figure 3 is only a high-level symbolic representation, and the exact detailed relationships can be more complicated.

5. ON CAUSAL INFERENCE AND ITS PLACE IN THE CAUSAL PRESCRIPTIVE ANALYTICS FRAMEWORK

Causal inference is an important tool within the causal prescriptive analytics framework, and the direct marketing example provided in section 2 illustrates how it can be utilized in Step 5 of the framework. We note, however, that causal inference is not always necessary to identify the relationship between X and Y.

How do you differentiate between contexts where causal inference is required versus not required? It is helpful to think of this question through the illustration in Figure 4. The outcome (Y) is a function of the decision variables (X) and some coefficients (C) that are often estimated from data. If there is a system response mechanism to the decision or intervention represented by the decision variables in the optimization problem X, then the relationship $X \rightarrow C$ needs to be taken into consideration when estimating Y, and causal inference is required. We outline a few examples next.

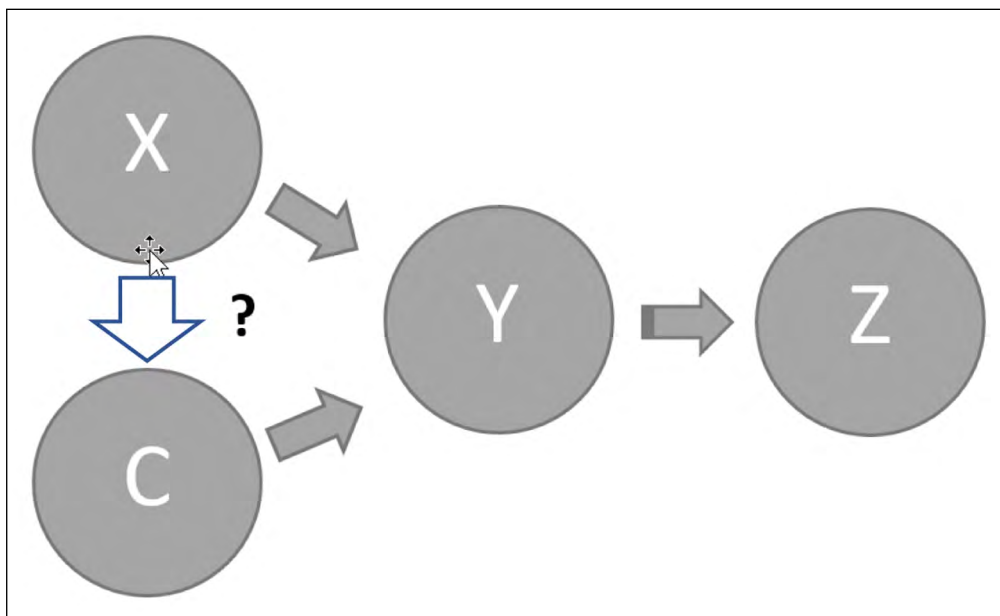


Figure 4 Causal inference in prescriptive analytics problem formulations.

5.1 CAUSAL INFERENCE *NOT* REQUIRED

Many important prescriptive analytics contexts do not require causal inference. This situation happens when there is no system response mechanism to a decision or intervention. Given the choice of a decision (i.e., a specific value of the decision variable), the immediate causal effect on the outcome variable is mathematically known (or straightforward to obtain). We will refer to Panel A of Table 1 for the examples below.

Vehicle routing. Because selecting an arc and adding it to a bus route is not really an intervention—it causally affects the total distance traveled (Y) and also the goal (Z) but does not affect the arc distance itself (C)—the standard vehicle routing problem does not require causal inference.

Workforce scheduling. The action of assigning an employee to a shift (X) does not impact the cost for staffing the employee to the shift (C), which is known in advance; so although the assignments affect the total cost (Y) causally, causal inference is not needed to estimate the relationship between X and Y.

Inventory management. Values such as annual usage rate and holding cost per year per item of inventory carried often need to be estimated using historical data (see Shapiro (2007), Pinedo (2009), or Daskin (2010)); however, given the known cost per order and holding cost per unit per year, the mathematical relationship between X (order quantity) and Y (order and holding costs) is typically known (assuming the company is a small player in the market), and causal inference is not required.

Portfolio construction. To maximize Z through Y, a key step is to predict the expected individual stock returns (C) using historical data or simulations; see, for example, Fabozzi et al. (2007). Once the expected individual stock returns are predicted, the future expected portfolio return can be calculated as a sum-product of the expected individual returns and the weights of the stocks in the portfolio. Although there is still estimation involved when it comes to determining C, the relationship between X, C, and Y is a mathematical formula that typically does not require causal inference because it is assumed that the portfolio is small relative to the market, so the decision to assign a particular weight to a stock does not affect the expected return on the stock.

5.2 CAUSAL INFERENCE REQUIRED

Causal inference is needed when the decision variable or intervention affects the function representing the relationship between X and Y and, by extension, the ultimate goal Z. Several categories of applications typically require causal inference (see Panel B of Table 1 for a technical summary):

1. Behavioral relationships. Human behaviors in response to individual-level stimuli or interventions are usually unknown and need to be estimated through causal inference.

Pricing. Estimating customer responses to pricing decisions typically requires causal inference, which is similar to the direct marketing problem in section 2.² This is because the decision variable price (X) affects sales volume (C) through an unknown price elasticity, thus the outcome sales revenue (Y), and ultimately the objective (Z), profitability. Common methodologies for estimating price elasticity include testing various price points through RCT, analyzing historical observational data through econometric time series analysis if there was price variation in historical data and RCT is not feasible, and survey-based conjoint or discrete choice analysis if in-market price changes are infeasible or difficult.

Customer retention. Because of limited resources, organizations have to select customers for retention efforts (an intervention), for example, outbound call programs with an incentive. The behavioral outcome for change in retention rate (C) and its relationship with the retention effort (X) is not known in advance and thus needs to be estimated through RCT by testing a combination of treatments, such as incentive, time/day of outreach, frequency, and channel, which is methodologically similar to the direct marketing example. If RCT is not available, causal inference techniques can be applied to nonexperimental data.

Employee acquisition. Since sales volume (C) may depend on available customer support and customer satisfaction, it is not a simple or known function of the number of staff (X) and may require causal inference to estimate; see Pessach et al. (2020). For example, one may utilize store-to-store variation as well as variation over time on the number of sales agents to develop a panel data analysis for assessing the impact on outcome metrics.

² Price testing could be included as part of direct marketing, as in many retail marketing programs.

2. Health care and medical examples. Similar to behavioral problems, human responses to health care services and medical treatments are typically unknown without utilizing causal inference.

Digital health. Causal inference techniques (e.g., RCT) are required to understand the effect of a message (X) on message-specific health outcomes (C). If multiple messages are eligible for each individual in a sequential order, the message-specific outcomes can be aggregated to the overall individual health outcome (Y), such as the number of emergency room visits, which is then translated to medical cost and summarized to the employer level (Z).

Personalized medicine. Medicine can be optimally assigned to appropriate patients (X) in order to maximize population health (Z) through individual treatment effectiveness (C); see Hamburg & Collins (2010) and Yong (2015). The effect of X on C is commonly measured via RCT. If multiple treatments are available, the overall effect at the individual level (Y) can be obtained through the estimates of C. For example, when the number of vaccines available in a country is limited, health officials have to decide who receives vaccination first in order to achieve a maximum protection for the whole population.

3. Policy examples. Government policies have an impact on individuals, organizations, and society as a whole, and such impact is usually assessed through observational data analysis, such as econometric methods, since RCT is often not feasible.

Health care policy. An example is a setting studied in Gai and Pachamanova (2019), where the impact of a policy (X) known as the Hospital Readmissions Reduction Program (HRRP) (as part of the Affordable Care Act (ACA)) is analyzed in an effort to reduce excess hospital readmissions ($Y = C$ in this case) and lower health costs (Z) while ensuring equitable treatment for vulnerable populations. Difference-in-difference was employed in the study to compare the pre-HRRP differences in readmission rates between treatment and control groups with their post-HRRP differences.

Economic policy. One of the most powerful economic decisions for central banks is setting an appropriate interest rate in order to stimulate the economy when it is weak or to prevent inflation when the economy is too strong. Setting it inappropriately can lead to an undesirable ripple effect. See Belongia and Ireland (2015) and Kiley and Roberts (2017) for examples of applying advanced econometric methods to measure impact on various metrics.

6. CONCLUSION

In this article, we introduced a causal prescriptive analytics framework that outlines the integration of multiple types of analytical techniques, including predictive analytics, machine learning, causal inference, and constrained optimization. These methodologies are from a variety of academic disciplines. We asserted that all prescriptive problems for optimal decision-making are causal by nature: in order to achieve a goal, we make a decision (or take an action) to cause a desirable outcome to happen. However, not all prescriptive problems require causal inference to uncover these relationships. We listed numerous examples where the framework can be applied and discussed several practical examples to illustrate the differentiation between problems that require causal inference and problems that do not. We also emphasized the importance of aligning the representation of causality with the ultimate goal. Our practical framework unifies decision-making problems in a common setting, facilitating the discovery of analytics opportunities and transitioning analytical decision-making from data science toward impactful decision science.

APPENDIX A: CONSTRAINED OPTIMIZATION FORMULATION OF THE MULTITREATMENT DIRECT MARKETING PROBLEM

This appendix describes the direct marketing problem mentioned in section 2 in more detail. Traditional response modeling based on conventional supervised learning aims at estimating the response rate, p_{ij} , for individual i to receive treatment j . Lo (2002; 2008); Siegel (2011; 2013); Kane, Lo, and Zheng (2014); Lo and Pachamanova (2015); and Haughton et al. (2023) have explained and demonstrated that such an approach is flawed, as scientific marketing measures “lift over control” by comparing the response outcome to a control group where treatment is

not given in order to causally assess whether a program is successful. To be consistent with this scientific measurement, uplift modeling is required to estimate “lift over control” as opposed to traditional modeling for estimating response rate only, which may capture customers who would naturally respond without receiving a treatment, resulting in inefficient targeting and potential waste of resources.

The treatment optimization problem can be formulated as a binary integer programming optimization model that can be solved by exact or heuristic methods. The objective function is to maximize incremental profitability due to direct marketing.

$$\text{Maximize } Z = rY = \sum_{i=1}^N \sum_{j=1}^M r \Delta \hat{p}_{ij} x_{ij} \quad (\text{A.1})$$

Subject to

$$\sum_{i=1}^N \sum_{j=1}^M c_{ij} x_{ij} \leq B, \quad \text{Budget Constraint}$$

$$\sum_{j=1}^M x_{ij} \leq 1, \quad \text{for } i = 1, \dots, N,$$

$$x_{ij} = 0 \text{ or } 1, \quad i = 1, \dots, N; \quad j = 1, \dots, M.$$

Here Z = incremental profit due to direct marketing; Y = incremental sales due to direct marketing; $\Delta \hat{p}_{ij}$ (represented by C in [Table 1](#)) = estimated lift value (treatment effect over no treatment) for individual i to receive treatment j ; r = revenue per sale (assumed constant here but can be relaxed); x_{ij} (decision variable) = 1 if treatment j is assigned to individual i and 0 otherwise; and c_{ij} = cost of promoting treatment j to individual i . (We provide this example to illustrate the main concepts in our framework. We note that that, in practice, the optimization problem can involve multiple treatments and products, and the causal relationships can be more complicated to estimate.)

In the above constrained optimization model, the constant r is irrelevant to the optimal solution and thus can be dropped (or assumed to be 1.0), and $\Delta \hat{p}_{ij}$ are the key input values that can be estimated by uplift modeling or CATE techniques based on historical RCT data; see [Lo & Pachamanova \(2015\)](#), [Pachamanova et al. \(2020\)](#), and [Haughton et al. \(2023: ch.6–7\)](#). The literature also describes a set of uplift modeling techniques for handling observational data through propensity score matching types of causal inference techniques; see [Athey & Imbens \(2015\)](#) and [Haughton et al. \(2023: ch.9\)](#).

To empirically illustrate the benefit of causal prescriptive analytics using uplift modeling over regular supervised learning for traditional response modeling in this application, that is, using estimated lift values $\Delta \hat{p}_{ij}$ as opposed to estimated response rates \hat{p}_{ij} in (A.1), we use online retail data for women’s and men’s merchandise from the Hillstrom data set, MineThatData (minethatdata.com). We follow the clustering-based heuristic optimization procedure described in [Lo and Pachamanova \(2015\)](#) and solve the optimization problem (A.1) to find the optimal treatment quantity for each customer segment.

The procedure from [Lo and Pachamanova \(2015\)](#) is outlined as follows:

1. Develop an uplift model for estimating lift in response rate using the separate model approach,³ which requires development of logistic regression models using the training data for treatment and control groups, respectively.
2. In the holdout data, compute the lift estimates for both men’s and women’s merchandise using the estimated uplift model by subtracting the estimated control response rate from the estimated treatment response rate at the individual level.
3. Perform a cluster analysis of individuals using the two lift estimates for men’s and women’s merchandise as input variables.
4. For each cluster in the holdout data, calculate the cluster-specific sample lift scores for both men’s and women’s merchandise by taking the difference between the sample mean response rate in the treatment group and the sample mean response rate in the control group.

³ The separate model approach is one of the uplift modeling techniques used for illustration here. See [Kane et al. \(2014\)](#) and [Athey and Imbens \(2015\)](#) for other techniques.

5. Apply the cluster solution to the new data⁴ for future campaigns.
6. Solve the integer program equivalent of (A.1) at the cluster level to maximize overall incremental value.

To evaluate the benefit of using uplift modeling over traditional response modeling, we repeat the above procedure using the estimated response rates \hat{p}_{ij} for men’s and women’s merchandise as input variables to cluster analysis in Step 3, as opposed to using the lift estimates, $\Delta \hat{p}_{ij}$, followed by applying the resulting cluster solution to the new data in Step 5 and determining the optimal solution to maximize overall value (instead of incremental value) in Step 6. We then compare the results of the two solutions using the objective function values in (A.1).

The results from the optimization using \hat{p}_{ij} and $\Delta \hat{p}_{ij}$ as objective function coefficients are shown in Table A.1a and Table A.1b, respectively.

Table A.1a Optimization using traditional response modeling.

CLUSTER	CLUSTER SIZE (IN NEW DATA)	MEN'S MER-CHANDISE TREATMENT RESPONSE RATE	WOMENS MERCHANDISE TREATMENT RESPONSE RATE	CONTROL RESPONSE RATE	MEN S MERCHANDISE LIFT IN RESPONSE	WOMEN'S MER-CHANDISE LIFT IN RESPONSE	DECISION VAR (TREATMENT QUANTITY) ON MEN'S	DECISION VAR (TREATMENT QUANTITY) ON WOMEN'S	TOTAL TREATMENT QUANTITY BY CLUSTER
1	3180	0.2549	0.2385	0.2617	-0.0068	-0.0232	3,180	-	3,180
2	40	0.1779	0.1477	0.1039	0.0741	0.0439	-	-	-
3	9110	0.3133	0.2425	0.1837	0.1296	0.0589	9,110	-	9,110
4	1090	0.5385	0.2273	0.2051	0.3333	0.0221	1,090	-	1,090
5	51300	0.1080	0.0793	0.0451	0.0628	0.0347	-	-	-
5	27950	0.2106	0.1588	0.1315	0.0791	0.0273	-	-	-
7	67170	0.1620	0.1440	0.0948	0.0672	0.0492	-	-	-
5	4220	0.3704	0.2817	0.2345	0.1359	0.0472	4,220	-	4,220
9	257590	0.2218	0.2116	0.1393	0.0826	0.0724	42,400	-	-
Total							obj value 5,597	-	5,597

CLUSTER	CLUSTER SIZE (IN NEW DATA)	MEN'S MER-CHANDISE TREATMENT RESPONSE RATE	WOMENS MER-CHANDISE TREATMENT RESPONSE RATE	CONTROL RESPONSE RATE	MEN S MERCHANDISE LIFT IN RESPONSE	WOMEN'S MER-CHANDISE LIFT IN RESPONSE	DECISION VAR (TREATMENT QUANTITY) ON MEN'S	DECISION VAR (TREATMENT QUANTITY) ON WOMEN'S	TOTAL TREATMENT QUANTITY BY CLUSTER
1	4,180	0.2333	0.0970	0.0746	0.1587	0.0224	4,180	-	4,180
2	5,650	0.2275	0.1568	0.1623	0.0652	-0.0055	-	-	-
3	60,220	0.1697	0.1668	0.1040	0.0658	0.0628	2,340	-	2,340
4	12,370	0.2854	0.2181	0.1563	0.1290	0.0618	12,370	-	12,370
5	8,940	0.1133	0.1221	0.0461	0.0672	0.0760	-	8,940	8,940
5	29,240	0.1626	0.1320	0.1107	0.0519	0.0213	-	-	-
7	28,070	0.2090	0.1475	0.1222	0.0868	0.0254	28,070	-	28,070
5	4,100	0.4194	0.2183	0.1944	0.2249	0.0239	4,100	-	4,100
9	37,060	0.1216	0.1071	0.0645	0.0572	0.0426	-	-	-
Total	189,850						obj value 5,773	680	6,453

Table A.1b Optimization using uplift modeling.

⁴ Following Lo and Pachamanova (2015), we assume the new data for future campaign is 10 times of the holdout data from the previous campaign.

The results in [Table A.1a](#) and [Table A.1b](#) are obtained by following exactly the same optimization procedure except for using \hat{p}_{ij} (noncausal) and $\Delta\hat{p}_{ij}$ (causal), respectively.⁵ The third, fourth, and fifth columns report the sample response rates for the men’s merchandise treatment group, women’s merchandise treatment group, and the control group, respectively, in the holdout data. The next two columns represent the lift estimates, which are the differences between men’s or women’s and the control group’s sample response rates by cluster. The two columns labeled “decision variables” are values from the optimal solution under each scenario. The last column is the sum of the previous two columns. The objective function values are listed at the lower right corner in both tables, which are the sums of incremental values from men’s and women’s merchandise. The percentage improvement in objective function value for using uplift modeling (causal) over traditional response modeling (noncausal) is $(6,453/5,597 - 1) \times 100\% = 15.3\%$. In [Table A.1a](#), to see the impact of using the treatment response rates for optimization as opposed to lift estimates, we observe the relatively high treatment response rates for cluster 1 resulting in an assigned quantity of 3,180 using the traditional approach, but this cluster has negative lift values when control response rate is considered, which would contribute negatively to the correct objective function value that is based on lift values.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Victor S. Y. Lo  orcid.org/0000-0003-3715-6722

Senior Vice President of Data Science and Artificial Intelligence, Workplace Investing, Fidelity Investments, Boston, MA 02210, US

Dessislava A. Pachamanova  orcid.org/0000-0002-1373-1553

Professor and Zwerling Family Endowed Term Chair, Babson College, Wellesley, MA 02457, US

REFERENCES

- Ali, SAM, Sorooshian, S and Kie, CJ.** 2016. Modelling for causal interrelationships by DEMATEL. *Contemporary Engineering Sciences*, 9(9): 403–412. DOI: <https://doi.org/10.12988/ces.2016.6214>
- Angrist, JD and Pischke, J-S.** 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton, NJ: Princeton University Press. DOI: <https://doi.org/10.1515/9781400829828>
- Antunes, CH, Alves, MJ and Climaco, J.** 2016. *Multiobjective Linear and Integer Programming*. Cham, Switzerland: Springer. DOI: https://doi.org/10.1007/978-3-319-28746-1_6
- Athey, S and Imbens, GW.** 2015. Machine learning methods for estimating heterogeneous causal effects. Working Paper No. 3350, Stanford Graduate School of Business.
- Attri, R, Dev, N and Sharma, V.** 2013. Interpretive structural modelling (ISM) approach: An overview. *Research Journal of Management Sciences*, 2(2): 3–8.
- Belongia, MT and Ireland, PN.** 2015. Interest rates and money in the measurement of monetary policy. *Journal of Business & Economic Statistics*, 33(2): 255–269. DOI: <https://doi.org/10.1080/07350015.2014.946132>
- Ben-Tal, A, El Ghaoui, L and Nemirovski, A.** 2009. *Robust optimization (Vol. 28)*. Princeton university press.
- Bertsimas, D, Delarue, A, Eger, W, Hanlon, J and Martin, S.** 2020. Bus routing optimization helps Boston public schools design better policies. *INFORMS Journal on Applied Analytics*, 50(1): 37–49. DOI: <https://doi.org/10.1287/inte.2019.1015>
- Bertsimas, D and Kallus, N.** 2020. From predictive to prescriptive analytics. *Management Science*, 66(3): 1025–1044. Available at <http://www.p2-analytics.com/papers/PredToPresc.pdf>. DOI: <https://doi.org/10.1287/mnsc.2018.3253>
- Bertsimas, D and Tsitsiklis, JN.** 1997. *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific.
- Bobriakov, I.** 2019. Data science vs. decision science. Medium, April 16, 2019. Available at <https://medium.com/@ibobriakov/data-science-vs-decision-science-infographic-7ad6e16698d>
- Bojinov, I, Chen, A and Liu, M.** 2020. The importance of being causal. *Harvard Data Science Review*, 2(3). <https://hdsr.mitpress.mit.edu/pub/wjth9tr/release/1>. DOI: <https://doi.org/10.1162/99608f92.3b87b6b0>

⁵ We set the number of clusters = 10 in both scenarios and combined insignificantly small clusters. The budget constraint is set at \$60,000, with cost per treatment = \$1, as outlined in Lo and Pachamanova (2015). In [Table A.1b](#), all metrics for cluster 2 are replaced with those for the overall holdout data due to its small size.

- Carpenter, SM, Menictas, M, Nahum-Shani, I, Wetter, DW and Murphy, SA.** 2020. Developments in mobile health just-in-time adaptive interventions for addiction science. *Current Addiction Reports*, 7: 280–290. DOI: <https://doi.org/10.1007/s40429-020-00322-y>
- Chauhan, A, Singh, A and Jharkharia, S.** 2018. An interpretive structural modeling (ISM) and decision-making trial and evaluation laboratory (DEMATEL) method approach for the analysis of barriers of waste recycling in India. *Journal of the Air & Waste Management Association*, 68(2): 100–110. DOI: <https://doi.org/10.1080/10962247.2016.1249441>
- Cornuejols, G and Tutuncu, R.** 2007. *Optimization Methods in Finance*. New York: Cambridge University Press.
- Daskin, MS.** 2010. *Service Science*. Hoboken, NJ: Wiley. DOI: <https://doi.org/10.1002/9780470877876>
- Davenport, TH.** 2013. Analytics 3.0. *Harvard Business Review*, 91(12): 64–72.
- de Langhe, B and Puntoni, S.** 2021. What leaders get wrong with data-driven decisions. *MIT Sloan Management Review: MIT's Journal of Management Research and Ideas*, 62(3): 10–13. Available at <http://hdl.handle.net/1765/134544>
- Delen, D.** 2020. *Prescriptive Analytics: The Final Frontier for Evidence-Based Management and Optimal Decision Making*. New York: Pearson.
- Dominici, F, Bargagli-Stoffi, FJ and Mealli, F.** 2021. From controlled to undisciplined data: Estimating causal effects in the era of data science using a potential outcome framework. *Harvard Data Science Review*. DOI: <https://doi.org/10.1162/99608f92.8102afed>
- Dunning, T.** 2016. *Natural Experiments in the Social Sciences: A Design-Based Approach*. 6th printing. Cambridge, UK: Cambridge University Press.
- Fabozzi, FJ, Kolm, PN, Pachamanova, DA and Focardi, SM.** 2007. *Robust Portfolio Optimization and Management*. Hoboken, NJ: Wiley. DOI: <https://doi.org/10.1002/9780470404324.hof003068>
- Feillet, D.** 2010. A tutorial on column generation and branch-and-price for vehicle routing problems. *4or*, 8(4): 407–424. DOI: <https://doi.org/10.1007/s10288-010-0130-z>
- Frazzetto, D, Nielsen, TD, Pedersen, TB and Siksny, L.** 2019. Prescriptive analytics: A survey of emerging trends and technologies. *VLDB Journal*, 28(4): 575–595. DOI: <https://doi.org/10.1007/s00778-019-00539-y>
- Freedman, DA.** 2009. *Statistical Models: Theory and Practice*. Rev. ed. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511815867>
- Gai, Y and Pachamanova, D.** 2019. Impact of the Medicare Hospital Readmissions Reduction Program on Vulnerable Populations. *BMC Health Services Research*, 19: 837. DOI: <https://doi.org/10.1186/s12913-019-4645-5>
- Gelman, A, Carlin, JB, Stern, HS and Rubin, DB.** 2013. *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: CRC Press. DOI: <https://doi.org/10.1201/b16018>
- Gelman, A, Hill, J and Vehtari, A.** 2021. *Regression and Other Stories*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/9781139161879>
- Glennerster, R and Takavarasha, K.** 2013. *Running Randomized Evaluations: A Practical Guide*. Princeton, NJ: Princeton University Press. DOI: <https://doi.org/10.2307/j.ctt4cgd52>
- Hagmayer, Y and Fernback, PM.** 2017. Causality in decision-making. In: Waldmann, MR (ed.), *The Oxford Handbook of Causal Reasoning*, 495–514. New York: Oxford University. DOI: <https://doi.org/10.1093/oxfordhb/9780199399550.013.27>
- Hagmayer, Y and Sloman, SA.** 2005. Causal models of decision making: Choice as intervention. In: *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society*, Stresa, Italy.
- Hamburg, MA and Collins, FS.** 2010. The path to personalized medicine. *New England Journal of Medicine*, 363(4): 301–304. DOI: <https://doi.org/10.1056/NEJMp1006304>
- Hastie, T, Tibshirani, R and Friedman, J.** 2009. *The Elements of Statistical Learning*. 2nd ed. New York: Springer. DOI: <https://doi.org/10.1007/978-0-387-84858-7>
- Haughton, D, Haughton, J and Lo, VSY.** 2023, expected. *Cause-and-Effect Business Analytics*. New York: CRC/Chapman & Hall.
- Hayes, AF.** 2013. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York: Guilford Press.
- Hernan, MA and Robins, JM.** 2020. *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC.
- Iacobucci, D.** 2008. *Mediation Analysis*. Thousand Oaks, CA: SAGE Publications.
- Illari, P and Russo, F.** 2014. *Causality: Philosophical Theory Meets Scientific Practice*. New York: Oxford University Press.
- Imbens, G and Rubin, D.** 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences: An Introduction*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139025751>
- Joyce, JM.** 2008. *The Foundations of Causal Decision Theory*. New York: Cambridge University Press.
- Kaliszewski, I, Miroforidis, J and Podkopaev, D.** 2016. *Multiple Criteria Decision Making by Multiobjective Optimization: A Toolbox*. Cham, Switzerland: Springer. DOI: <https://doi.org/10.1007/978-3-319-32756-3>

- Kane, K, Lo, VSY, and Zheng, J.** 2014. Mining for the truly responsive customers and prospects using true-lift modeling: Comparison of new and existing methods. *Journal of Marketing Analytics*, 2(4): 218–238. DOI: <https://doi.org/10.1057/jma.2014.18>
- Keeney, RL.** 1996. *Value-Focused Thinking: A Path of Creative Decisionmaking*. Cambridge, MA: Harvard University Press. DOI: <https://doi.org/10.2307/j.ctv322v4g7>
- Kiley, MT and Roberts, JM.** 2017. Monetary policy in a low interest rate world. Finance and Economics Discussion Series 2017-080. Washington, DC: Board of Governors of the Federal Reserve System. DOI: <https://doi.org/10.17016/FEDS.2017.080>
- Kochenderfer, MJ.** 2015. *Decision Making under Uncertainty*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/10187.001.0001>
- Koller, D and Friedman, N.** 2009. *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, MA: MIT Press.
- Koole, G.** 2013. *Call Center Optimization*. Amsterdam: MG Books.
- LaRiviere, J, McAfee, P, Rao, J, Narayanan, VK and Sun, W.** 2016. Where predictive analytics is having the biggest impact. *Harvard Business Review*, May 25, 2016. <https://hbr.org/2016/05/where-predictive-analytics-is-having-the-biggest-impact>
- Leigh, A.** 2018. *Randomistas: How Radical Researchers Are Changing Our World*. New Haven, CT: Yale University Press. DOI: <https://doi.org/10.12987/9780300240115>
- Lo, VSY.** 2002. The true lift model—A novel data mining approach to response modeling in database marketing.” *SIGKDD Explorations*, 4(2): 78–86. https://www.researchgate.net/publication/220520042_The_True_Lift_Model_-_A_Novel_Data_Mining_Approach_to_Response_Modeling_in_Database_Marketing. DOI: <https://doi.org/10.1145/772862.772872>
- Lo, VSY.** 2008. New opportunities in marketing data mining. In: Wang, J (ed.), *Encyclopedia of Data Warehousing and Mining*, 2nd ed. Hershey, PA: Idea Group Publishing. DOI: <https://doi.org/10.4018/978-1-59904-951-9.ch177>
- Lo, VSY.** 2020. Top 10 essential data science topics to real-world application from the industry perspectives. *Harvard Data Science Review*, 2(3) (Summer). <https://hdsr.mitpress.mit.edu/pub/diub13so/release/3>
- Lo, VSY and Pachamanova, D.** 2015. A practical approach to treatment optimization while accounting for estimation risk. *Journal of Marketing Analytics*, 3(2): 79–95. DOI: <https://doi.org/10.1057/jma.2015.5>
- Markowitz, H.** 1952. Portfolio selection. *Journal of Finance* 7(1): 77–91. DOI: <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- Mills, TC and Markellos, RN.** 2008. *The Econometric Modelling of Financial Time Series*. 3rd ed. Cambridge, UK: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511817380>
- Morgan, SL and Winship, C.** 2015. *Counterfactuals and Causal Inference*. 2nd ed. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781107587991>
- Pachamanova, D, Lo, VSY and Gulpinar, N.** 2020. Uncertainty representation and risk management in direct segmented market. *Journal of Marketing Management*, 36(1–2): 149–175. DOI: <https://doi.org/10.1080/0267257X.2019.1707265>
- Papadimitriou, CH and Steiglitz, K.** 1998. *Combinational Optimization: Algorithms and Complexity*. Mineola, NY: Dover.
- Pearl, J.** 2000. *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.
- Pearl, J.** 2012. The do-calculus revisited. In: De Freitas, N and Murphy, K (eds.), *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, 4–11. Corvallis, OR: AUAI Press. https://ftp.cs.ucla.edu/pub/stat_ser/r402.pdf
- Pearl, J, Glymour, M and Jewell, NP.** 2016. *Causal Inference in Statistics: A Primer*. Hoboken, NJ: Wiley.
- Pearl, J and MacKenzie, D.** 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.
- Pessach, D, Singer, G, Avrahami, D, Chalutz Ben-Gal, H, Shmueli, E and Ben-Gal, I.** 2020. Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming. *Decision Support Systems*, 134: 113290. DOI: <https://doi.org/10.1016/j.dss.2020.113290>
- Peters, J, Janzing, D and Scholkopf, B.** 2017. *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press: Cambridge, MA.
- Pinedo, ML.** 2009. *Planning and Scheduling in Manufacturing and Services*. 2nd ed. New York: Springer. DOI: <https://doi.org/10.1007/978-1-4419-0910-7>
- Poornima, S and Pushpalatha, M.** 2020. A survey on various applications of prescriptive analytics. *International Journal of Intelligent Networks*, 1: 76–84. DOI: <https://doi.org/10.1016/j.ijin.2020.07.001>
- Powell, WB.** 2011. *Approximate Dynamic Programming*. 2nd ed. Hoboken, NJ: Wiley. DOI: <https://doi.org/10.1002/9780470400531.eorms0043>
- Reichardt, CS.** 2019. *Quasi-Experimentation: A Guide to Design and Analysis*. New York: Guilford Press.
- Rose, R.** 2016. Defining analytics: A conceptual framework. *ORMS Today*, 43(3). DOI: <https://doi.org/10.1287/orms.2016.03.12>

- Rosenbaum, PR.** 2002. *Observational Studies*. 2nd ed. New York: Springer. DOI: <https://doi.org/10.1007/978-1-4757-3692-2>
- Rosenbaum, PR.** 2010. *Design of Observational Studies*. New York: Springer. DOI: <https://doi.org/10.1007/978-1-4419-1213-8>
- Rosenbaum, PR.** 2019. *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.
- Routing Challenge.** 2021. <https://routingchallenge.mit.edu/about-the-challenge/>
- Rubin, DB.** 2006. *Matched Sampling for Causal Effects*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511810725>
- Salsburg, D.** 2001. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. New York: Holt Paperbacks.
- Scutari, M** and **Denis J-B.** 2015. *Bayesian Networks: With Examples in R*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Shapiro, JF.** 2007. *Modeling the Supply Chain*. 2nd ed. Belmont, MA: Thomson Brooks/Cole.
- Siegel, E.** 2011. Uplift modeling: Predictive analytics can't optimize marketing decisions without it. Prediction Impact white paper sponsored by Pitney Bowes Business Insight.
- Siegel, E.** 2013. *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*. Hoboken, NJ: Wiley.
- Sloman, SA.** 2005. *Causal Models: How People Think about the World and Its Alternatives*. New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195183115.001.0001>
- Sloman, SA** and **Hagmayer, Y.** 2006. The causal psycho-logic of choice. *TRENDS in Cognitive Sciences*, 10(9): 407–412. DOI: <https://doi.org/10.1016/j.tics.2006.07.001>
- Sorooshian, S, Tavana, M** and **Ribeiro-Navarrete, S.** 2023. From classical interpretive structural modeling to total interpretive structural modeling and beyond: A half-century of business research. *Journal of Business Research*, 157: 113642. DOI: <https://doi.org/10.1016/j.jbusres.2022.113642>
- Spirtes, P, Glymour, C** and **Scheines, R.** 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/1754.001.0001>
- Sugiyama, M.** 2015. *Statistical Reinforcement Learning: Modern Machine Learning Approaches*. Boca Raton, FL: CRC Press. DOI: <https://doi.org/10.1201/b18188>
- Sutton, RS** and **Barto, AG.** 2018. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, MA: MIT Press.
- Theodoridis, S.** 2015. *Machine Learning: A Bayesian and Optimization Perspective*. London: Academic Press. DOI: <https://doi.org/10.1016/B978-0-12-801522-3.00012-4>
- Thomke, SH.** 2020. *Experimentation Works: The Surprising Power of Business Experiments*. Boston, MA: Harvard Business Review Press.
- Vanderweele, TJ.** 2015. *Explanation in Causal Inference: Methods for Mediation and Interaction*. New York: Oxford University Press. DOI: <https://doi.org/10.1093/ije/dyw277>
- Wallace, SW** and **Ziemba, WT.** (eds.) 2005. *Applications of Stochastic Programming*. Philadelphia, PA: Society for Industrial and Applied Mathematics and the Mathematical Programming Society.
- Williams, HP.** 2003. *Model Building in Mathematical Programming*. 4th ed. West Sussex, UK: Wiley.
- Yong, FH.** 2015. Quantitative methods for stratified medicine. PhD dissertation, Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University.

TO CITE THIS ARTICLE:

Lo, VSY and Pachamanova, DA. 2023. From Meaningful Data Science to Impactful Decisions: The Importance of Being Causally Prescriptive. *Data Science Journal*, 22: 8, pp. 1–18. DOI: <https://doi.org/10.5334/dsj-2023-008>

Submitted: 18 February 2022

Accepted: 06 March 2023

Published: 25 April 2023

COPYRIGHT:

© 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.