Long-Term Policy Prediction with Causal Graphs: Insights from Healthy Grocery Shopping

Nur Kaynar
Samuel Curtis Johnson Graduate School of Management, Cornell University, nur.kaynar@cornell.edu

Dmitry Mitrofanov
Carroll School of Management, Boston College, dmitry.mitrofanov@bc.edu.

Policymakers frequently implement short-term interventions (e.g., temporary healthy food subsidies) to achieve long-term outcomes (e.g., lasting healthy eating habits). While the delay in observing these outcomes presents challenges for immediate evaluation of effectiveness of these interventions, we can observe short-term variables to get early insights into the long-term effects. However, identifying which of these short-term variables will reliably indicate long-term changes remains challenging. To address this challenge, we introduce a causal structure learning algorithm to uncover the causal relations among policy intervention, short-term variables, and long-term outcome. Our algorithm combines short-term experimental and long-term observational data and returns a graph in which causal relations are specified by directed edges. By mapping out the underlying causal relations, we determine (*i*) which short-term variables are affected by the policy change and (*ii*) which of these variables subsequently impact the long-term outcome. We develop an identification strategy to consistently estimate the long-term treatment effect of a policy without waiting for the long-term outcome to materialize. By selecting short-term outcomes (surrogates) and covariates to control for, based on the causal relationships discovered in the graph, we derive a novel closed-form expression to compute the long-term treatment effect. Our results demonstrate that the proposed method enables earlier estimation of long-term effects, offering policymakers a valuable tool to refine their policy implementation strategies. Additionally, our findings offer novel empirical insights into how government agencies can effectively stimulate health-conscious grocery shopping. Our findings show that targeting specific population segments for healthy product subsidies, especially those with historically low consumption of healthy products, can produce lasting positive effects. In particular, younger consumers in urban areas demonstrate a greater likelihood of maintaining healthier eating habits even after the subsidies end. Overall, our paper provides policymakers with valuable insights and practical tools to enhance public healthy product subsidy strategies. By allowing them to evaluate long-term impacts through short-term randomized controlled trials, the framework we present helps refine policy design and optimize effectiveness.

Key words: causal inference; surrogates; long-term policy effects; machine learning; healthy food subsidy; natural experiment; empirical findings

1. Introduction

Understanding the long-term consequences of policy interventions is essential in ensuring a policy's effectiveness and sustainability. For example, assesing the lasting impacts of government subsidies for healthy eating on dietary habits can be critical in ensuring that these subsidies justify their

costs and result in enduring positive outcomes (Hinnosaar 2023). Similarly, before making job training programs widely available, identifying their long-term effects on labor market outcomes can be essential for achieving the desired results (Hotz et al. 2006). However, focusing on long-term outcomes makes it difficult for policymakers to assess the effectiveness of interventions as tracking units over a long period of time is prohibitively costly or sometimes infeasible.

Similar challenges also arise in other fields beyond policymaking. In digital technology companies, experimentation methods such as A/B tests are fundamental for assessing the effects of marketing or operational decisions. Although A/B tests provide precise short-term insights, they often fall short of capturing long-term impacts due to the fast-paced nature of business operations, where waiting for extended periods is impractical. This limitation underscores the difficulty in making timely decisions based on long-term effects (Gupta et al. 2019). In fact, there are plenty of examples where the treatment effects measured in short-term experiments do not accurately reflect the actual long-term impact. For instance, Dekimpe et al. (1998) assess the immediate and prolonged effects of price promotions on different brands spanning across multiple product categories. Their findings indicate that while price promotions can have a positive short-term impact on revenue, their long-term effects can be mixed and may not always benefit the company. Moreover, in reality, the accuracy of long-term predictions may also depend on the degree of pricing personalization (Jagabathula et al. 2022) and price competition (Caro and Martínez-de Albéniz 2012), how strategic the decision makers are (Chen et al. 2019, Esenduran et al. 2020), and the dynamic nature of the pricing decisions (Farias and Van Roy 2010, Caro and Gallien 2012).

To bridge the gap between short-term results and long-term impacts, the seminal paper of Athey et al. (2019) propose a framework that is based on combining short-term experimental data with long-term historical sample through surrogate variables. A surrogate variable is an intermediate outcome that can be measured in a shorter period of time and mediates the effects of interventions. While initial studies focused on using a single surrogate (Prentice 1989), Athey et al. (2019) propose a surrogate index methodology that is based on combining multiple surrogates to improve the accuracy of long-term effect estimation. Specifically, this approach involves using experimental data where the treatment and surrogates are observed, but the long-term outcome is not. This is complemented by historical data that includes both the surrogates and the long-term outcome but does not include the treatment. Then the average treatment effect on the long-term outcome is identified from the combination of these two samples without the need to wait for the outcome to be realized. Athey et al. (2019) has led to numerous subsequent studies. Yang et al. (2023) extends this idea for policy optimization, introducing an experimental framework designed to directly optimize targeting policies for long-term customer retention and revenues. Other papers address the confounding factors when using observational datasets to infer the relationship between short-term surrogates and

longer-term outcomes (Athey et al. 2020, Imbens et al. 2022). In a similar spirit, Huang et al. (2023) study the problem of estimating the "long-term effects of long-term treatments" which is different from estimating the "carryover effects of short-term treatments". Battocchi et al. (2021) develop a methodology that extends beyond traditional surrogate-based approaches by accommodating continuous treatments and treatment policies with serial correlation. Anderer et al. (2022) propose a Bayesian adaptive clinical trial design that uses data from surrogate and true outcomes to improve decision-making. All existing research assumes that the surrogates mediating the treatment effect are known a priori. However, identifying surrogates can be challenging because it requires precise knowledge of the underlying causal pathways through which the treatment affects the outcome.

In this work, we complement recent surrogacy approaches by proposing a causal structure learning algorithm that combines experimental and observational data to uncover the *underlying causal structure* among treatment, long-term outcome, and other short-term variables. Causal structure learning, unlike causal inference, focuses on identifying the presence or absence of causal relationships between variables using the framework of graphical causal models (Pearl 2000, Spirtes et al. 2000). The proposed causal structure learning algorithm returns a directed acyclic graph (DAG) that visually represents these causal relationships. By mapping out these relationships, the causal graph helps determine (*i*) which short-term variables are influenced by the treatment and (*ii*) which of those variables, in turn, affect the long-term outcome. This is particularly valuable for decision-makers, as it reveals relationships that may be unknown or extend beyond intuition and expert opinion, enabling a data-driven method to identify surrogate variables that mediate the treatment effect. Using the inferred causal graph, we then derive a closed-form expression for the average long-term treatment effect, proving that it can be consistently estimated from the available experimental and historical data under the standard assumptions in the surrogacy literature. We demonstrate the performance of the proposed framework in estimating the lasting effects of temporary healthy food subsidies using data from the U.S. Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). Our empirical results demonstrate that the proposed method enables earlier estimation of long-term effects, potentially reducing experimental costs and allowing for timely decision-making. In this regard, our paper effectively bridges the field of operations with graphical causal models, econometric methods, and machine learning. Moreover, at a higher level, our paper is related to the recent research in the operations management field studying experimental design (Johari et al. 2022, Farias et al. 2022, Bojinov et al. 2023, Xiong et al. 2024), causal inference (Ho et al. 2017, Kallus and Zhou 2021, Farias et al. 2021, Wang et al. 2022, Zhou et al. 2023, Ye et al. 2023, Singal and Michailidis 2024, Eberhardt et al. 2024), and graph-based models (Lu and Van Mieghem 2009, Bayati et al. 2018).

Our paper makes a significant contribution not only in terms of methodology but also in advancing the understanding of the ongoing challenge of incentivizing healthy consumption through subsidies. Despite nearly two decades since the World Health Organization[1] first highlighted this issue, there remains a lack of clarity on the long-term effectiveness of the mechanisms designed to positively impact the customers' diet. While previous research has generally suggested that healthy product subsidies have only a short-lived impact on consumer behavior (Hinnosaar 2023), our study presents new evidence. We find that targeting specific segments of the population – particularly those with a history of low healthy product consumption – can yield positive, lasting effects. Furthermore, within this segment, younger consumers living in urban areas show a greater likelihood of maintaining healthier consumption patterns even after the subsidies are removed. Our research provides preliminary but compelling evidence for the need for more nuanced policies that consider the diverse characteristics and geographic differences among consumers. Rather than a broad, 'one-size-fits-all' approach, effective healthy product subsidy related interventions may require tailoring strategies to specific demographic and regional contexts to optimize long-term outcomes in healthy product consumption.

Overall, our study first highlights the practical use of our methodological framework, which integrates surrogate models and causal discovery techniques, to forecast the long-term effectiveness of healthy product subsidies. In addition to introducing this innovative approach, we offer empirical evidence that pinpoints key dimensions where the policy can achieve meaningful success. These findings not only confirm the validity of our framework but also offer practical insights into specific areas where the policy adjustments can be most effective. Furthermore, our paper equips policymakers with the tools to refine public subsidy strategies by enabling them to assess long-term impacts through field experiments in a significantly shorter timeframe. By providing a data-driven approach to evaluate and adjust the policy, we believe this work offers critical insights for improving the design and implementation of healthy product subsidies, ensuring they are more effective and targeted in promoting lasting behavioral change.

1.1. Contributions

The major contributions of our study can be outlined as follows:

- *Causal structure learning algorithm that integrates experimental and observational samples.* First, we present the COMB-PC algorithm, a causal structure learning algorithm that integrates experimental and observational samples to uncover the underlying causal relationships among the variables of interest (see details in §4). This algorithm consists of two major stages.

[1] World Health Organization (2004) Global strategy on diet, physical activity and health. http://apps.who.int/gb/ebwha/ pdf$_files/WHA57/A57_R17-en.pdf(accessedMay2012)$.

In the first stage of the COMB-PC algorithm, we focus on determining the *skeleton* of the causal graph, which is an undirected graph depicting potential causal links without orienting the direction of these links. Then, the second stage of the COMB-PC algorithm orients the undirected edges present in the skeleton identified in the first stage. The COMB-PC algorithm, through its phases of skeleton discovery and edge orientation, remains in spirit with the essence of the PC algorithm (Spirtes et al. 2000). However, it introduces tailored adaptations to suit our surrogacy framework, particularly with modifications that enable integrating short-term experimental data with long-term observational data. Notably, this is the first study that uses causal structure learning within a surrogacy framework, bridging two previously distinct yet fundamentally connected research areas.[2]

- *Long-term treatment effect identification strategy.* Building on the discovered causal structure, we develop a novel non-parametric identification strategy for the average long-term treatment effect using short-term experimental data and long-term observational data (see details in §5). More specifically, we use a two-step algorithmic framework for the graphs learned by the COMB-PC algorithm. The first step focuses on finding a valid set of surrogate variables that mediate the treatment effect for a given graph. The second step involves selecting an appropriate adjustment set to control for confounders that may influence both the identified surrogates and the long-term outcome. Using the identified surrogates and control variables, we derive a closed-form formula that identifies the long-term treatment effect using both the experimental and observational samples.
- *Validation of the proposed frameworks using synthetic data.* We numerically evaluate the accuracy of both the COMB-PC algorithm and the proposed identification strategy using synthetic data across varying graph densities (see details in §6). Our results indicate that the accuracy of causal structure learning is high for sparser graphs but decreases as graph density increases. We find that this reduced accuracy in causal structure learning for denser graphs then affects the accuracy of the treatment effect estimations, underscoring the importance of correctly identifying underlying causal graphs in obtaining reliable treatment effect estimations.
- *Case study: analyzing the long-term effects of subsidies on healthy food products.* In our real-world case study, we demonstrate the performance of the proposed framework in estimating the long-term effects of temporary healthy food subsidies within the context of grocery shopping behavior. Specifically, we focus on assessing the impact during the second and third years after the subsidy program ends. We use short-term variables observed within the first year

[2] Imbens (2020) identifies the surrogacy setting as an area that could benefit from the use of graphical causal models to clarify underlying assumptions. In this work, we take this one step forward and propose a framework to *learn* surrogate variables using graphical causal models.

post-subsidy, such as changes in purchasing patterns and consumer responsiveness to discounts as potential surrogates. We first discover the underlying causal relations among the intervention, short-term variables, and the long-term outcome using the COMB-PC algorithm. Then using the discovered graph, we identify the surrogate variables and control variables to use in estimating the long-term effect. Our results demonstrate that we can accurately estimate the impact of these subsidies on the second and third years through the identified surrogates by using only one year of data.

- *Health-conscious grocery shopping: managerial/societal insights.* Importantly, in addition to the methodological contribution, our paper makes a substantial empirical contribution to the literature on health-conscious grocery shopping. As government agencies increasingly recognize the importance of promoting healthy product choices for public health, our research identifies a specific subset of individuals who could benefit more significantly from healthy product subsidies, though such subsidies might not be universally effective as shown by Hinnosaar (2023). Our novel findings suggest that it might be worth targeting healthy product subsidies towards individuals with limited exposure to healthy products, as they are likely to experience a more sustainable impact, ensuring permanence in healthy product consumption and thus a more efficient use of taxpayer funds. This targeting strategy might be grounded in two key mechanisms: the *informational effect*, where subsidies can educate consumers about the benefits of healthy eating, and the *consideration set expansion effect*, making consumers likely to try and continue purchasing healthy products after initial financial incentives.[3] Conversely, individuals already frequently purchasing healthy products may only respond to these subsidies in the short term. Thus, while broad subsidy policies like the reform of the U.S. special supplemental nutrition program for women, infants, and children may not alter long-term purchasing behaviors across all households, targeting those less exposed to healthy options could create lasting changes in their consumption patterns.
- *The long-term impact of healthy product subsidies: heterogeneous treatment effect.* We also find that younger consumers, along with those residing in high-density, urban areas, are more likely to maintain healthy consumption habits even after subsidies are discontinued. This suggests two key points: first, younger consumers appear more inclined to develop lasting healthy eating habits when nudged by short-term subsidies. Second, the increased exposure to a broader range of healthy food options and the convenience of accessing these products in urban areas likely contribute to sustained behavioral changes. These factors emphasize the importance of targeted policies that consider both demographic and geographic characteristics to enhance the long-term effectiveness of these public health interventions.

[3] Although verifying these mechanisms empirically is challenging using our current data.

The remainder of the paper is organized as follows. Section §2 reviews foundational concepts in graphical causal models, which may be skipped by readers already familiar with the topic. Section §3 formally defines the problem and discusses the key assumptions. Section §4 presents the causal structure learning algorithm. Section §5 provides an identification framework to estimate the long-term treatment effect using causal graphs. Sections §6 and §7 examine the performance of the proposed framework based on the synthetic experiment and real-world dataset, respectively. Additionally, we conduct a comprehensive empirical study in Section §7, providing significant policy insights into the effectiveness of healthy product subsidies for targeted consumer groups. Section §8 concludes.

2. Graphical Causal Models

Graphical causal models (Pearl, 2000; Spirtes et al., 2000) use directed graphs to represent causal relationships among multiple variables. Let $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ be a directed graph, where $\mathbf{V}$ denotes a set of nodes and $\mathbf{E}$ represents the set of edges, such that $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$. In this graphical representation, each edge in $\mathcal{G}$ signifies a direct causal connection between the corresponding nodes.

We next define some graphical preliminaries that are used throughout the paper. Any two nodes X and Y are called adjacent if there is an edge $X \to Y$ or $X \leftarrow Y$ in the corresponding graph $\mathcal{G}$. The *parents* and *children* of a node X represent its direct causes and effects, respectively, in the graph $\mathcal{G}$. We say that a node Y is a *collider* on a path if its adjacent edges point into Y, i.e., $\to Y \leftarrow$. A *noncollider* on a path is a node Y that is either a *mediator* ($\to Y \to$) or a *common cause* ($\leftarrow Y \to$). In Figure 1(a), node Y is a collider on path $X \to Y \leftarrow Z$, in Figure 1(b) node Y is a mediator on path $X \to Y \to Z$, and in Figure 1(c) node Y is a common cause on path $X \leftarrow Y \to Z$. A *v-structure*, also known as an *unshielded collider*, is a specific configuration of nodes in a graph. A v-structure consists of two parent nodes directing edges towards a common child node, without an edge between the parents, forming a "V" shape. Any node that is connected to node X by a directed path is called a *descendant* of X, while any node connected to X by a directed path is an *ancestor* of X. We refer to the skeleton of a graph $\mathcal{G}$ as the undirected graph obtained by replacing directed edges in $\mathcal{G}$ with undirected edges. For instance, while the edge orientations vary among the graphs in Figure 1, their underlying skeletons are identical.

We define a *path* between two nodes X and Y as a sequence of nodes that starts with X, ends with Y, and where each consecutive pair of nodes in the sequence is connected by an edge in the graph. In this context, the direction of the edges does not matter; as long as there is a connection between any two consecutive nodes in the sequence, it forms a valid path. In other words, although the edge directions are considered, they do not impose any constraints on constructing a path between nodes. A *directed path* from node X to node Y is a path in which all edges point towards node Y. Then, a *directed acyclic graph (DAG)* is a directed graph without cycles.

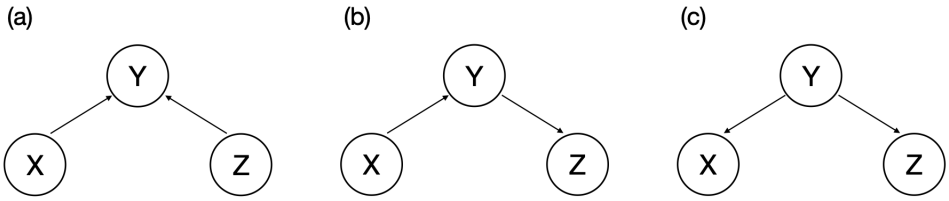

Figure 1 Illustrations of Collider, Mediator, and Common Cause Structures.

2.1. Causal Structure Learning

Causal modeling involves associating a probability distribution, denoted as $P_{\mathcal{G}}(\mathbf{V})$, with a graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, which represents the causal relationships among the variables or nodes in the set $\mathbf{V}$. The underlying assumption is that the distribution $P_{\mathcal{G}}(\mathbf{V})$ is generated by the graph structure in a way that allows factorization: $P_{\mathcal{G}}(\mathbf{V}) = \prod_{X \in V} P_{\mathcal{G}}(X|Pa(X))$, where $Pa(X)$ represents the parents of node X in $\mathcal{G}$ (Spirtes and Zhang 2016, Eberhardt 2017). In this context, the terms "node" and "variable" can be used interchangeably, as they correspond to the graphical structure and the probability distribution, respectively.

The two key assumptions that bridge the observed data and the causal structure are stated below.

Assumption 1 **(Causal Markov)**. *Each variable $X \in \mathbf{V}$ in a graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ is probabilistically independent of its non-descendants given its parents.*

Assumption 2 **(Faithfulness)**. *The only independences present in the probability distribution are those that are implied by the graph structure through the causal Markov conditions.*

Assumption 3 **(Causal Sufficiency)**. *For any pair of variables in $\mathbf{V}$, all common causes of those variables are also contained within $\mathbf{V}$.*

The first assumption, the causal Markov condition, permits us to transition from the causal graph to the observed probabilistic independencies. Conversely, the faithfulness condition enables us to deduce the structure of the causal graph from observed data independencies. Lastly, the causal sufficiency assumption ensures that all common causes of any pair of variables in the set $\mathbf{V}$ are also contained within $\mathbf{V}$, thereby excluding the existence of any hidden or unobserved confounders.

Due to the close relationship between the causal structure and the resulting data distribution, many algorithms for causal structure learning leverage the identifiable independence structure in the data to make inferences about the underlying causal relationships. A key concept essential for this inference is *d-separation* (Geiger et al. 1990), often considered as the graphical equivalent of probabilistic independence. It is based on the notion of a *blocked path*:

DEFINITION 1 (BLOCKED PATHS). A path between nodes X and Y is considered *unblocked* with respect to a set of nodes $\mathbf{C}$ if every collider Z on the path is in $\mathbf{C}$ or has a descendant in $\mathbf{C}$, and no other nodes on the path are in $\mathbf{C}$. If these conditions do not hold, the path is considered *blocked* with respect to $\mathbf{C}$ (Pearl 2000).

We can now introduce the concept of d-separation.

DEFINITION 2 (D-SEPARATION). Two nodes X and Y are said to be *d-separated* with respect to a conditioning set $\mathbf{C}$ (denoted as $i \perp j|\mathbf{C}$) if all paths between them are blocked. Conversely, if there exists at least one unblocked path between X and Y given $\mathbf{C}$, they are considered *d-connected* (denoted as $i \not\perp j|\mathbf{C}$) (Pearl 2000).

REMARK 1. Under the causal Markov and faithfulness conditions, a (conditional) independence in $P_{\mathcal{G}}(\mathbf{V})$ is present if and only if there is a corresponding (conditional) d-separation in DAG $\mathcal{G}$ (Pearl 2000).

Remark 1 highlights the correspondence between (conditional) independence in $P_{\mathcal{G}}(\mathbf{V})$ and (conditional) d-separation in DAG $\mathcal{G}$ under the causal Markov and faithfulness conditions. This correspondence serves as the fundamental framework for a wide range of causal structure learning methods, providing a means to leverage observed independence patterns in data to infer underlying causal relations.

To illustrate the notion of blocked paths and their connection to the principles outlined in Remark 1, we refer to Figure 1. In Figure 1(a), we see the path $X \to Y \leftarrow Z$ where Y serves as a collider. Given the conditioning set $\mathbf{C} = \emptyset$, this path is considered 'blocked' since the collider, Y, is not included in $\mathbf{C}$. The path $X \to Y \leftarrow Z$ is the only path between X and Z within this figure. Since it is blocked, we establish that X and Z are d-separated, signifying a lack of information flow between the two nodes. By invoking Remark 1, we anticipate that X and Z to be marginally probabilistically independent, in line with their status of being d-separated within the graph. On the other hand, when $\mathbf{C} = Y$, the path $X \to Y \leftarrow Z$ transitions to being 'unblocked'. This is because Y, being the only collider on this path, is now included within the conditioning set $\mathbf{C}$. Hence by Remark 1, we expect that X and Z to be probabilistically dependent with respect to conditioning set $\mathbf{C} = Y$. Conversely, in Figure 1 (b) and (c), the paths $X \to Y \to Z$ and $X \leftarrow Y \to Z$ respectively are 'unblocked' when the conditioning set is $\mathbf{C} = \emptyset$. This is due to the absence of colliders on these paths and the fact that Y, acting as a noncollider, is not included in the conditioning set $\mathbf{C}$. Hence, we expect that X and Z to be marginally dependent. However, conditioning on $\mathbf{C} = Y$ blocks these paths and X and Z become marginally independent with respect to conditioning set $\mathbf{C} = Y$.

Despite different causal relationships in Figure 1(b) and 1(c), they imply the same independence relations. In contrast, the independence relations in Figure 1(a), i.e., $X \perp\!\!\!\perp Z$ and $X \not\perp\!\!\!\perp Z \mid Y$ uniquely identifies Y as a collider. These differences underscore the fact that colliders leave distinct signatures

on conditional independence patterns. However, mediator and common causes can result in identical patterns in conditional independence relations. This demonstrates that the observable data cannot uniquely identify the underlying causal graph, as both mediators and common causes can generate identical patterns. This idea is captured by the concept of *Markov equivalence class*. Two graphs that have the same independence structure are said to be *Markov equivalent*. This means that the independence relations represented by each graph in the equivalence class are identical, even though the causal relationships they represent are not. Two DAGs are in the same Markov equivalence class if and only if they have the same skeleton and the same v-structures (Verma and Pearl 1990).

Causal structure learning methodologies have historically been focused on analyzing individual data sets. Nevertheless, in recent years, there has been a notable shift towards broadening the traditional approach, driven by the rapid surge in available data sets from both observational and experimental sources. As a result, there is a growing body of work centered on causal structure learning over multiple observational or experimental datasets. In the literature on causal structure learning methods, this problem has been addressed in two main ways. One group of papers focuses on developing algorithms that combine observational data measuring overlapping variables (Tillman et al. 2008, Triantafillou et al. 2010, Tillman and Spirtes 2011, Claassen and Heskes 2010). Another line of research has concentrated on integrating multiple experimental data on *identical* set of variables (Cooper and Yoo 2013, Tong and Koller 2001, Mooij et al. 2020, Zhang et al. 2017) and *non-identical/overlapping* set of variables (Triantafillou and Tsamardinos 2015, Huang et al. 2020). In this paper, we contribute to this stream of work by combining short-term experimental data with long-term observational data over *non-identical/overlapping* variables to learn the underlying causal structure. While various methods could potentially be adapted to this setting, we propose an extension to the PC-algorithm (Spirtes et al. 2000), selected for its widespread recognition and intuitive nature.

2.2. Identification of Average Treatment Effect using Causal Graphs

The previous section discusses how to discover causal graphs, which are essential for learning presence or absence of causal relationships among variables. However, understanding the structure of underlying causal relations is only the first step. Many policy decisions are also interested in quantifying the causal effects of specific interventions. From the joint distribution of two variables, W and Y, we can derive the conditional probability $P(Y \mid W)$ using observational data, which tells us the probability of Y given that W takes on a specific value w, i.e., $P(Y \mid W = w)$. However, what we often need for policy decisions is the causal effect of setting W to a specific value w, represented as $P(Y \mid \mathrm{do}(W = w))$. This notation distinguishes the intervention $\mathrm{do}(W = w)$ from mere observation.

While we can calculate $P(Y \mid W = w)$ directly from the joint distribution of W and Y in an observational data, the challenge lies in determining whether this observational data, combined with the underlying causal structure, allows us to infer the causal effect $P(Y \mid \text{do}(W = w))$.

With this notation, given a graph $\mathcal{G}$, the average effect of a binary treatment can be defined as:

$$\tau_{\mathcal{G}} = E[Y|do(W = 1)] - E[Y|do(W = 0)].^{4} \tag{1}$$

Identification of the average treatment effect $\tau_{\mathcal{G}}$ from observational data is challenging primarily due to the presence of confounders that can induce spurious correlations between the treatment W and the outcome Y. For example, let's consider the causal graph depicted in Figure 2. In this graph, there are three paths between W and Y: $W \to X_1 \to X_2 \leftarrow X_3 \to Y$, $W \to X_4 \to Y$, and $W \leftarrow X_5 \to X_6 \to Y$. By Definition 1, the path $W \to X_1 \to X_2 \leftarrow X_3 \to Y$ is *blocked* as the variable X_2 is a collider and both $W \to X_4 \to Y$ and $W \leftarrow X_5 \to X_6 \to Y$ are *unblocked* as there are no colliders on these paths. However, only the path $W \to X_4 \to Y$ represents a causal effect of treatment W on outcome Y. The path $W \leftarrow X_5 \to X_6 \to Y$, in contrast, reflects confounding. The association observed between W and Y along this path is not due to a causal mechanism from W to Y, but is a byproduct of their mutual associations through X_5 and X_6. Thus, without controlling for the confounders X_5 or X_6, the observed correlation in the data between W and Y will reflect both the actual causal effect from W to Y and the spurious correlation introduced by the path via X_5 and X_6. To accurately estimate the average treatment effect τ, it is vital to adjust for these confounders and thereby isolate the causal effect of W on Y.

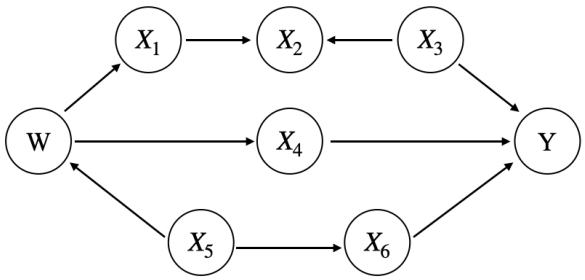

Figure 2 A figure to illustrate the back-door criterion.

As demonstrated in the example above, adjusting for relevant covariates is crucial to accurately estimate causal effects from observational data. This method, known as covariate adjustment, helps isolate the true causal relationship between the treatment and the outcome by controlling for confounders that could otherwise bias the results (Pearl 1995, Shpitser et al. 2010). One of the most

[4] In the potential outcomes framework, $do(W = 1)$ and $do(W = 0)$ correspond to the notation $Y(1)$ and $Y(0)$, respectively. Therefore, $E[Y|do(W = 1)] - E[Y|do(W = 0)] = E[Y(1)] - E[Y(0)]$. For more details, see Rubin (1974) and Imbens and Rubin (2015).

well-known methods for this is Pearl's back-door adjustment. When estimating the effect of W on Y, a *back-door path* is defined as any unblocked path connecting W to Y that begins with an arrow pointing towards W. Such paths introduce potential confounding by offering a non-causal route for information flow between W and Y. Pearl (2000) introduced the *back-door criterion* as a method for identifying sets of variables that, when conditioned on, can block the backdoor paths and allow for the estimation of causal effects from observational data. The back-door criterion provides a graphical test to determine a valid adjustment set to estimate the causal effect from observational data.

DEFINITION 3. **Back-Door Criterion (Pearl 2000).**[5] Let $\mathbf{W}, \mathbf{Y}$ and $\mathbf{Z}$ be pairwise disjoint sets of vertices in a DAG $\mathcal{G}$. $\mathbf{Z}$ satisfies the *back-door criterion* relative to $\mathbf{W}, \mathbf{Y}$ in $\mathcal{G}$ if

(i) no node in $\mathbf{Z}$ is a descendant of any node in $\mathbf{W}$, and

(ii) for every $W \in \mathbf{W}$, the set $\mathbf{Z} \cup \mathbf{W} \setminus \{W\}$ blocks every back-door path from W to any member of $\mathbf{Y} \in \mathcal{G}$.

Theorem 1 ***Back-Door Criterion (Pearl 2000).*** *If a set of variables* $\mathbf{Z}$ *satisfies the back-door criterion relative to* $(\mathbf{W}, \mathbf{Y})$ *in a DAG* $\mathcal{G}$*, then the causal effect of* $\mathbf{W}$ *on* $\mathbf{Y}$ *is identifiable and is given by the formula*

$$P(\mathbf{Y} = \mathbf{y} | do(\mathbf{W} = \mathbf{w})) = \sum_{\mathbf{z} \in \mathcal{Z}} P(\mathbf{Y} = \mathbf{y} | \mathbf{W} = \mathbf{w}, \mathbf{Z} = \mathbf{z}) P(\mathbf{Z} = \mathbf{z}) \tag{2}$$

where $\mathcal{Z}$ *is the support of* $\mathbf{Z}$*.*

Corollary 1.1 ***ATE with Back-Door Criterion.*** *Let* $\mathbf{W}$ *be a binary treatment variable. If a set of variables* $\mathbf{Z}$ *satisfies the back-door criterion relative to* $(\mathbf{W}, \mathbf{Y})$ *in a DAG* $\mathcal{G}$*, then the ATE* $\tau_{\mathcal{G}}$ *of* W *on* Y *is identifiable and is given by the formula*

$$\tau_{\mathcal{G}} = \sum_{y \in \mathcal{Y}} \sum_{z \in \mathcal{Z}} \mathbf{y} P(\mathbf{Y} = \mathbf{y} | \mathbf{W} = \mathbf{1}, \mathbf{Z} = \mathbf{z}) P(\mathbf{Z} = \mathbf{z}) - \sum_{\mathbf{y} \in \mathcal{Y}} \sum_{\mathbf{z} \in \mathcal{Z}} y P(\mathbf{Y} = \mathbf{y} | \mathbf{W} = \mathbf{0}, \mathbf{Z} = \mathbf{z}) P(\mathbf{Z} = \mathbf{z}) \tag{3}$$

where $\mathcal{Y}$ *and* $\mathcal{Z}$ *are the supports of* $\mathbf{Y}$ *and* $\mathbf{Z}$*, respectively.*

This corollary demonstrates that if we have access to control variables satisfying the back-door criterion, we can accurately estimate the average treatment effect. This is achieved by calculating the expected outcomes conditioned on both the treatment and covariate values, and then weighting these outcomes by the probability of observing the covariate values. However, the back-door adjustment criterion is not complete (Pearl 2000). Thus, there are causal graphs where the back-door criterion

[5] The definition is adapted from Pearl's to explicitly consider cases where W and Y are sets of vertices, in accordance with the set-based analysis framework in Maathuis and Colombo (2015).

does not identify a valid adjustment set, but adjusting for different covariate sets can still accurately estimate the causal effect.

There is a vast amount of research focused on finding necessary and sufficient graphical criteria for the selection of adjustment sets to accurately estimate causal effects. Shpitser et al. (2010) extend the back-door criterion and provide a necessary and sufficient graphical criterion for adjustment in DAGs. Other research has focused on constructing adjustment sets that are valid for more general graph classes, including those with unobserved confounders or structures that represent Markov equivalence classes (van der Zander et al. 2014, Maathuis and Colombo 2015, Perkovi et al. 2018). However, these methods are designed for single-sample settings and are not directly applicable when combining multiple datasets. Our work aligns with a recent stream of research in causal graphical models that combines multiple datasets collected under heterogeneous conditions, such as different populations and various sampling methods, with the possibility of sampling biases (Bareinboim and Pearl 2012, Lee et al. 2020, Jung et al. 2024). However, these studies *do not* consider the scenario of short-term experimental data with missing long-term outcomes combined with long-term historical data. We contribute to this literature by algorithmically selecting surrogates and covariates to derive a *novel* closed-form expression for the long-term treatment effect that is *guaranteed* to be computed from the available samples.

3. Problem Setup

To begin with, let us consider a setting with two samples: an experimental sample (denoted as E) and an observational sample (denoted as O). The experimental sample includes N_E observations and comprises the treatment variable W along with some covariates $\mathbf{X}$. However, it does not include the primary outcome Y, which is only observed after a significant delay. The observational sample includes N_O observations and includes both the primary outcome Y and the covariates $\mathbf{X}$, but it does not include the treatment variable W. Let $\mathbf{V^E}$ represent the variables in the experimental sample, such that $\mathbf{V^E} = \{W\} \cup \mathbf{X}$. Similarly, let $\mathbf{V^O}$ represent the variables in the observational sample, defined as $\mathbf{V^O} = \{Y\} \cup \mathbf{X}$. The complete set of variables, encompassing both samples, is stored in $\mathbf{V}$, where $\mathbf{V} = \{W\} \cup \mathbf{X} \cup \{Y\}$. We use indicator $D \in \{E, O\}$ to denote the considered sample [6]

Our goal is twofold. First, we focus on learning the true underlying graph $\mathcal{G}^*$ that represents the causal relations among variables $\mathbf{V}$. Then, we estimate the long-term treatment effect based on the population from which the experimental sample is drawn:

$$\tau^* = E[Y|D = E, do(W = 1)] - E[Y|D = E, do(W = 0)].$$

[6] In this paper, we focus on a binary treatment variable W and discrete outcome Y, with $\mathbf{X}$ representing a set of discrete variables. However, our results can be extended to the continuous setting.

The main challenge in learning the underlying graph and estimating the treatment effect τ^* stems from the absence of the long-term outcome Y in the experimental sample. To address this issue, we complement the experimental sample with the observational sample. First, we propose a causal structure learning algorithm that is tailored for this two-sample setting to learn the underlying causal structure that governs the relations among variables in $\mathbf{V}$. This enables the identification of surrogate variables that mediate the treatment's impact on long-term outcomes. Second, we develop an identification strategy for the average long-term treatment effect using the discovered graph and surrogate variables.

3.1. Assumptions to Combine Experimental and Observational Data

In this section, we present the assumptions on the structures of the underlying graph, experimental data and observational data.

The first assumption formally states that we consider a standard randomized experiment, with subjects being randomly assigned to treatment and control groups.

ASSUMPTION 4. ***(Randomized Treatment Assignment).*** *The treatment W is randomly assigned, implying that the true underlying graph $\mathcal{G}^*$ does not include any incoming edges to the treatment node W.*

This assumption ensures that any observed differences in outcomes can be attributed to the treatment itself, rather than to underlying differences among the subjects. The graphical representation of the treatment node W without any incoming edges represents the absence of direct influences or confounding factors affecting the treatment variable. Additionally, this assumption implies strong ignorability or unconfoundedness in the potential outcomes framework.

Since our experimental sample lacks long-term outcomes, we introduce an assumption on the true causal structure over variables $\mathbf{V}$ to be able to estimate long-term treatment effect using an observational sample:

ASSUMPTION 5. ***(Structure of Causal Graph).*** *Let graph $\mathcal{G}^*$ be the true causal structure over the variables $\mathbf{V}$ in the experimental sample. We assume that*

(i) $\mathcal{G}^*$ does not include a direct edge between treatment W and the long-term outcome Y,

(ii) $\mathcal{G}^*$ does not include any outgoing edges from the long-term outcome Y to any other variable in $\mathbf{V}$.

Assumption 5(i) states that there is no direct edge between the treatment and the long-term outcome in the true causal structure over variables $\mathbf{V}$. Instead, the effect of the treatment is mediated entirely through the set of covariates $\mathbf{X}$, signifying that any influence the treatment has on the outcome must pass through these intermediate variables. Assumption 5(ii) states that Y is not a

cause of any other variables in $\mathbf{V}$. We believe this assumption is not overly restrictive as we are considering a setting where Y is observed after a long delay. Given that Y is observed later, it is unlikely that it has any causal effect on variables observed prior to it.

Finally, we introduce a standard assumption required for inferring long-term treatment effects by integrating short-term experimental data with observational data.

ASSUMPTION 6. ***(Comparability of Samples).*** *For any subset of variables $\mathbf{C} \subseteq \mathbf{V} \setminus \{W, Y\}$, the conditional distribution of Y given $\mathbf{C}$ remains the same across both observational and experimental samples:*

$$P(Y = y | D = E, \mathbf{C} = \mathbf{c}) = P(Y = y | D = O, \mathbf{C} = \mathbf{c}), \ \forall y \in \mathscr{Y}, \mathbf{c} \in \mathscr{C}, \mathbf{C} \subseteq \mathbf{V} \setminus \{W, Y\},$$

where $\mathscr{Y}$ is the support of Y and $\mathscr{C}$ is the support of $\mathbf{C}$.

This assumption ensures that, after controlling for a subset of variables $\mathbf{C}$, the conditional distribution of the long-term outcome Y remains consistent across both experimental and observational samples. Comparability of samples is a standard assumption in the surrogacy literature (Athey et al. 2019, Imbens et al. 2022, Yang et al. 2023, Huang et al. 2023). Similar assumptions regarding the use of causal estimates from one population to infer effects in another based on pre-treatment variable distributions have been utilized in prior research (Hotz et al. 2005, Hernán and VanderWeele 2011, Pearl and Bareinboim 2014). It should be noted that while this assumption is presented in its most restrictive form here, it can be relaxed somewhat after identifying the appropriate surrogates and control variables in §5.

4. Causal Structure Learning with Experimental and Observational Data

In this section, we introduce the COMB-PC algorithm, a causal structure learning algorithm that combines experimental sample E with observational sample O to learn the underlying causal relations among variables $\mathbf{V}$. We begin by defining the possible conditioning sets for the pairs of variables in both experimental and observational samples. For variables $V_i, V_j \in \mathbf{V}^E$, we store all possible conditioning sets in $\mathbf{C}^E_{V_i V_j} = \{\mathbf{C} \mid \mathbf{C} \subseteq \mathbf{V^E} \setminus \{V_i, V_j\}\}$. Similarly, for variables $V_i, V_j \in \mathbf{V}^O$, we store all possible conditioning sets in $\mathbf{C}^O_{V_i V_j} = \{\mathbf{C} \mid \mathbf{C} \subseteq \mathbf{V^O} \setminus \{V_i, V_j\}\}$.

In the first phase of the COMB-PC algorithm, we focus on identifying the *skeleton* of the causal graph. A skeleton is an undirected graph that represents potential causal relations among variables without specifying their directionality. This phase begins with the initialization of a complete undirected graph $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U})$ where $\mathbf{U}$ consists of all possible undirected edges between variables in $\mathbf{V}$ except the edge between the treatment W and the outcome Y. We intentionally exclude the edge between the treatment variable W and the long-term outcome Y as established in Assumption 5 (*ii*).

Algorithm 1: COMB-PC - Phase 1 (Skeleton Discovery)

Input: $\mathbf{C}^{E}_{V_i,V_j}\ \forall V_i, V_j \in \mathbf{V}^{E}, \mathbf{C}^{O}_{V_i,V_j}\ \forall V_i, V_j \in \mathbf{V}^{O}$
Output: $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U}), \mathcal{N}(V_i)$ for $V_i \in \mathbf{V}$.
Initialization: $\mathbf{U} = \{(V_i — V_j) \mid \forall\ V_i, V_j \in \mathbf{V^E}\} \cup \{(V_i — Y) \mid \forall\ V_i \in \mathbf{V^O}\}$,
$\mathcal{N}(W) = \mathbf{V} \setminus \{W, Y\}, \mathcal{N}(V_i) = \mathbf{V} \setminus \{V_i\}, \forall V_i \in \mathbf{V^E} \setminus \{W\}, \mathcal{N}(Y) = \mathbf{V} \setminus \{W, Y\}$,
$SepSet_{V_iV_j} = \emptyset, \forall V_i, V_j \in \mathbf{V}^{E}, \ell = 0$.

1. **Experimental sample:**
   for a pair $V_i, V_j \in \mathbf{V^E}$ where $V_i \in \mathcal{N}(V_j)$:
      for $\mathbf{C} \in \mathbf{C}^{E}_{V_iV_j}$ where $|\mathbf{C}| = \ell$:
         if $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}$:
            Update $SepSet_{V_iV_j} \leftarrow SepSet_{V_iV_j} \cup \{\mathbf{C}\}$ and $\mathbf{U} \leftarrow \mathbf{U} \setminus \{(V_i — V_j)\}$.
            Update $\mathcal{N}(V_i) \leftarrow \mathcal{N}(V_i) \setminus \{V_j\}$ and $\mathcal{N}(V_j) \leftarrow \mathcal{N}(V_j) \setminus \{V_i\}$.
            Break.
2. **Long-term outcome integration:**
   for $V_i \in \mathbf{V}^{O} \setminus Y$ where $V_i \in \mathcal{N}(Y)$:
      for $\mathbf{C} \in \mathbf{C}^{O}_{V_iY}$ with $|\mathbf{C}| = \ell$:
         if $V_i \perp\!\!\!\perp_O Y \mid \mathbf{C}$:
            Update $\mathbf{U} \leftarrow \mathbf{U} \setminus \{(V_i — Y)\}$, $\mathcal{N}(Y) \leftarrow \mathcal{N}(Y) \setminus \{V_i\}$ and $\mathcal{N}(V_i) \leftarrow \mathcal{N}(V_i) \setminus \{Y\}$.
            Break.
3. **if** $\ell < |\mathbf{V}| - 3$:
      Update $\ell = \ell + 1$. Go to Step 1.
4. Return $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U}), \mathcal{N}(V_i)\ \forall V_i \in \mathbf{V}^{E}, \mathcal{N}(Y), SepSet_{V_iV_j}\ \forall V_i, V_j \in \mathbf{V}^{E}$.

This way we can ensure that W does not have a direct causal impact on Y, instead, the effect of W on Y is mediated through other variables. Note that $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}$ and $V_i \perp\!\!\!\perp_O V_j \mid \mathbf{C}$ denote the probabilistic independence of V_i and V_j with respect to the conditioning set $\mathbf{C}$ in the experimental (E) and observational (O) samples, respectively. For every pair of variables $V_i, V_j \in \mathbf{V}^{E}$, we check if V_i and V_j are independent given a specific conditioning set $\mathbf{C} \in \mathbf{C}^{E}_{V_iV_j}$. If a conditioning set makes V_i and V_j independent, we remove the undirected edge between V_i and V_j from the graph $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U})$. We then test the independence of Y and each V_i for all V_i in $\mathbf{V^O} \setminus Y$. If V_i and Y are found to be independent given a particular conditioning set $\mathbf{C} \in \mathbf{C}^{O}_{V_iY}$, we then remove the undirected edge between V_i and Y from the graph $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U})$.

The second phase of the COMB-PC algorithm focuses on orienting the undirected edges within the skeleton returned in the first phase. The inputs for this phase are the skeleton with undirected edges, i.e., $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U})$, the neighborhood sets for each variable, i.e., $\mathcal{N}(V_i)$ for each $V_i \in \mathbf{V}$, and the separation set $SepSet_{V_iV_j}$ for each pair $V_i, V_j \in \mathbf{V}^{E}$, which stores the conditioning sets that found to make V_i and V_j independent in the previous phase. We proceed by examining conditional independence relations between variable triplets $V_i, V_j, V_k \in \mathbf{V}^{E}$ such that $V_j \neq W$ is a neighbor of V_i and V_k, but V_i is not a neighbor of V_k. By Definitions 1 and 2, if V_i and V_k are probabilistically dependent with respect to conditioning set $\mathbf{C} = \{V_j\}$, i.e., $V_j \notin SepSet_{V_iV_k}$, then V_j must be a collider between V_i and V_k. Hence, we orient the edge $V_i — V_j$ as $V_i \rightarrow V_j$ and the $V_j — V_k$ as $V_j \leftarrow V_k$. Additionally, for each

Algorithm 2: COMB-PC - Phase 2 (Edge Orientation)

Input: $\mathcal{G}_1 = (\mathbf{V}, \mathbf{U}), \mathcal{N}(V_i)\ \forall V_i \in \mathbf{V}^E, \mathcal{N}(Y), SepSet_{V_iV_j}\ \forall V_i, V_j \in \mathbf{V}^E$.
Output: $\mathbf{G}$.
Initialization: $\mathbf{M} = \mathbf{U}$, $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$.
1. **Experimental sample:**
   for $V_i, V_j, V_k \in \mathbf{V}^E$ where $V_i \in \mathcal{N}(V_j), V_j \in \mathcal{N}(V_k)$ and $V_i \notin \mathcal{N}(V_k)$:
      if $V_j \notin SepSet_{V_iV_k}$ and $V_j \neq W$:
         Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(V_i \rightarrow V_j), (V_j \leftarrow V_k)\}\big) \setminus \{(V_i — V_j), (V_j — V_k)\}$.
   Treatment randomization:
   for $V_i \in \mathbf{V}^E$ where $V_i \in \mathcal{N}(W)$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(W \rightarrow V_i)\}\big) \setminus \{(W — V_i)\}$.
2. **Long-term outcome integration:**
   for $V_i \in \mathbf{V}^O$ where $V_i \in \mathcal{N}(Y)$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(V_i \rightarrow Y)\}\big) \setminus \{(V_i — Y)\}$.
3. **Meek rules:**
(R1) **for** $V_i, V_j, V_k \in \mathbf{V}$ where $V_i \notin \mathcal{N}(V_k)$ and $\{(V_i \rightarrow V_j), (V_j — V_k)\} \subseteq \mathbf{M}$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(V_j \rightarrow V_k)\}\big) \setminus \{(V_j — V_k)\}$.
(R2) **for** $V_i, V_j, V_k \in \mathbf{V}$ where $\{(V_i \rightarrow V_j), (V_j \rightarrow V_k), (V_i — V_k)\} \subseteq \mathbf{M}$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(V_i \rightarrow V_k)\}\big) \setminus \{(V_i — V_k)\}$.
(R3) **for** $V_i, V_j, V_k, V_l \in \mathbf{V}$ where $V_j \notin \mathcal{N}(V_l)$ & $\{(V_i — V_j), (V_j \rightarrow V_k), (V_i — V_l), (V_l \rightarrow V_k), (V_i — V_k)\} \subseteq \mathbf{M}$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup \{(V_i \rightarrow V_k)\}\big) \setminus \{(V_i — V_k)\}$.
(R4) **for** $V_i, V_j, V_k, V_l \in \mathbf{V}$ where $V_i \notin \mathcal{N}(V_k)$ & $\{(V_i \rightarrow V_j), (V_j \rightarrow V_k), (V_i — V_l), (V_l — V_k), (V_j — V_l)\} \subseteq \mathbf{M}$:
      Update $\mathbf{M} \leftarrow \big(\mathbf{M} \cup (\{V_l \rightarrow V_k)\}\big) \setminus \{(V_l — V_k)\}$.
4. Identify all DAGs within the Markov equivalence class characterized by $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$ and store them in $\mathbf{G}$.
5. Return $\mathbf{G}$.

variable V_i in the experimental sample that is a neighbor of the treatment variable W, the edge $W — V_i$ is oriented as $W \rightarrow V_i$. This specific orientation is based on the randomized treatment assignment stated in Assumption 4, which implies that there are no incoming edges to the treatment variable W (Pearl 2000). In the long-term outcome integration step, for each variable V_i that is found to be a neighbor of the long-term outcome Y in the observational sample, the edge $V_i — Y$ is oriented as $V_i \rightarrow Y$. This orientation is based on Assumption 5(ii) that the long-term outcome Y does not affect the covariates $\mathbf{X}$. The algorithm then employs Meek's rules for additional edge orientation, ensuring a consistent and acyclic graph structure over the variables. These rules are illustrated in Appendix A (Meek 1995). Note that $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$ can include mixed edges, i.e., both directed and undirected edges, and represents a Markov equivalence class. In $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$, directed edges represent causal directions that the algorithm successfully identified whereas the undirected edges represent causal relations where directionality is not conclusively identified, allowing for different DAGs within the same Markov equivalence class to include edges with different directions while remaining consistent with the observed independencies. In the last step of phase 2, we identify all DAGs that are in the equivalence class represented by $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$ and store them in $\mathbf{G}$. Figure 3 illustrates the phases

of the COMB-PC algorithm. Note that the true causal graph is uniquely identified in this example; however, unique identification is not always guaranteed as it was mentioned above.

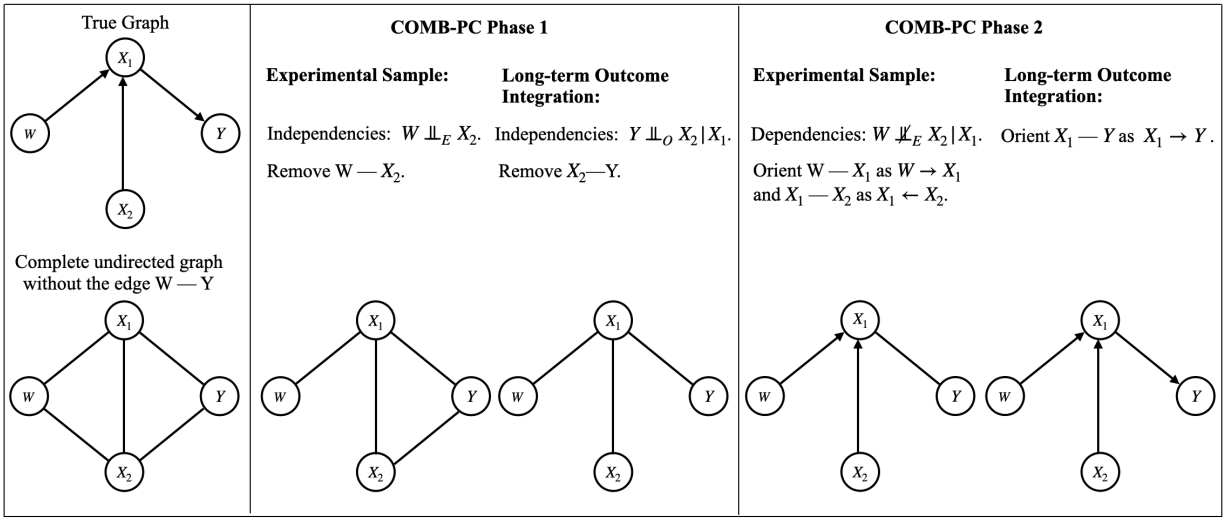

Figure 3 **Illustration of how the COMB-PC algorithm works.**

We assume that we have access to the results of all possible independence tests over a given set of variables and that the test results correctly describe an underlying ground truth DAG $\mathcal{G}^*$:

ASSUMPTION 7 **(Complete oracle)**. *Let $\mathcal{G}^*$ be the true data-generating graph. For all $V_i, V_j \in \mathbf{V}^E$ and $\mathbf{C} \in \mathbf{C}^E_{V_i V_j}$, $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}$ if and only if V_i and V_j are d-seperated with respect to $\mathbf{C}$ in graph $\mathcal{G}^*$. Furthermore, For all $V_i \in \mathbf{V}^O \setminus Y$ and $\mathbf{C} \in \mathbf{C}^O_{V_i Y}$, $V_i \perp\!\!\!\perp_O Y \mid \mathbf{C}$ if and only if V_i and Y are d-seperated with respect to $\mathbf{C}$ in graph $\mathcal{G}^*$.*

We can now state the main result of this section.

Theorem 2 *Let $\mathcal{G}^*$ be the true underlying DAG and $\mathbf{G}$ store the set of DAGs returned by the COMB-PC algorithm. Under Assumptions 1–7, we have $\mathcal{G}^* \in \mathbf{G}$ and $\mathcal{G}^*$ is Markov equivalent to any graph $\mathcal{G} \in \mathbf{G}$.*

Theorem 2 confirms the asymptotic correctness of the COMB-PC algorithm under appropriate assumptions. Assumption 7 allows us to separate the discovery task, handled by the COMB-PC algorithm, from the statistical inference of the conditional independence tests. The assumption also describes the (conditional) independence/dependence relations that would be obtained in the large-sample limit, which we use to prove the asymptotic correctness of our algorithm.

The COMB-PC algorithm, in both its skeleton discovery and edge orientation phases, closely mirrors the traditional PC algorithm (Spirtes et al. 2000), but with some modifications to combine

experimental and observational samples. In the skeleton discovery phase, similar to the traditional PC algorithm, the COMB-PC algorithm begins with an almost complete undirected graph and iteratively refines it by testing for conditional independence and removing edges accordingly. A key difference in the initial setup is the intentional exclusion of an edge between the treatment variable W and the outcome Y in the initial graph. This decision is pivotal in ensuring the validity of the surrogacy setting, where the causal effect of W on Y is mediated through other variables. Furthermore, since the treatment W and the long-term outcome Y are not jointly observed in either sample, we cannot directly test for independence between these variables. Therefore, excluding the edge between the treatment variable W and the outcome Y in the initial graph naturally aligns with this data limitation. Also, we ensure that the conditioning sets, $\mathbf{C}^{E}_{V_i V_j}$ for experimental samples and $\mathbf{C}^{O}_{V_i V_j}$ for observational samples, do not include both the treatment variable W and the outcome Y together, in line with the sample restrictions. Lastly, the COMB-PC algorithm diverges from the traditional PC method by orienting edges to account for the specific setting we consider where the treatment is randomized and the outcome is observed after a long delay. While the ingredients of the COMB-PC algorithm are not entirely new, its added value lies in extending the PC algorithm, which is the most commonly used causal structure learning algorithm, to address the unique challenges of the two-sample surrogacy framework we study. This extension effectively bridges two previously distinct yet fundamentally connected research areas.

We thus far established a framework to learn causal graphs that are consistent with the true underlying causal relations, using short-term experimental data and long-term historical data. Next, we propose a *novel* framework to estimate the long-term treatment effect by leveraging these graphs.

5. Long-term Treatment Effect Identification with Surrogates

In this section, we develop a long-term treatment effect identification strategy using the graphs in $\mathbf{G}$ returned by the COMB-PC algorithm. Our strategy is based on the algorithmic selection of valid surrogates and backdoor control variables for each graph $\mathcal{G} \in \mathbf{G}$ (refer to §2.2 for details on backdoor control variables). At a high level, our identification strategy extends Pearl's backdoor criterion to address the unique challenges of the two-sample framework we study. While the long-term treatment effect is always identifiable within the two-sample framework under the previously discussed assumptions in §3.1, the method of identification may vary depending on the graph. However, the proposed algorithms for selecting surrogates and backdoor covariates ensure a complete identification strategy for all graphs suitable to the surrogacy setting, unlike the backdoor criterion, which may fail to identify causal effects in some graphs (Pearl 2000).

We start by formally defining a valid set of surrogates.

Definition 4 (Surrogate Variables). A set $\mathbf{S} \subseteq \mathbf{V} \setminus \{W, Y\}$ is a *valid surrogate set* for the effect of treatment W on the long-term outcome Y in a DAG $\mathcal{G}$ if (1) $\mathbf{S}$ *blocks* all directed paths from W to Y in $\mathcal{G}$ and (2) each $S \in \mathbf{S}$ lies on a directed path from W to Y in $\mathcal{G}$.

Given a graph $\mathcal{G}$, a valid set of surrogates can be identified by examining all directed paths in the graph. Figure 4 illustrates this definition. In the DAG depicted in Figure 4, there are two directed paths from W to Y: $W \to X_3 \to X_4 \to Y$ and $W \to X_5 \to X_4 \to Y$. By Definition 4, several combinations of variables could serve as valid sets of surrogates. For example, the sets $\{X_3, X_5\}$, $\{X_3, X_4\}$, and $\{X_4\}$ are all valid. Each of these sets blocks all directed paths from W to Y and the graph includes a directed path from W to Y through the variables in these sets. However, the set $\{X_3\}$ is not valid because it does not block the path that goes via X_5 and X_4. Similarly, $\{X_1, X_4\}$ is also invalid, as it includes X_1, which is not on any directed path from W to Y.

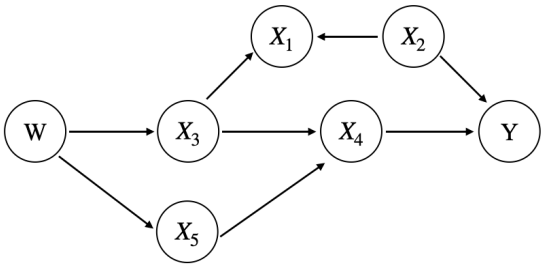

Figure 4 A figure to illustrate the surrogate definition.

The following proposition characterizes the effect of W on a set of surrogates within the experimental sample over a graph $\mathcal{G}$.

Proposition 1. *Suppose Assumption 4 holds. Let* $\mathbf{S}_\mathcal{G}$ *be a valid surrogate set for graph* $\mathcal{G}$*, as defined in Definition 4. Then the causal effect of* W *on the surrogate set* $\mathbf{S}_\mathcal{G}$ *over graph* $\mathcal{G}$ *within the experimental sample is given by*

$$P(\mathbf{S}_\mathcal{G} = \mathbf{s}_\mathcal{G} | D = E, do(W = w)) = P(\mathbf{S}_\mathcal{G} = \mathbf{s}_\mathcal{G} | D = E, W = w). \tag{4}$$

Proposition 1 indicates that in the experimental sample, the interventional distribution $P(\mathbf{S}_\mathcal{G} = \mathbf{s} | D = E, do(W = w))$ is equivalent to the observed distribution $P(\mathbf{S}_\mathcal{G} = \mathbf{s} | D = E, W = w)$. The underlying intuition of Proposition 1 arises from the randomization of the treatment W, as stated in Assumption 4. The randomization ensures that the treatment is independent of any confounders, allowing for the direct estimation of its causal effect on the set of surrogates $\mathbf{S}_\mathcal{G}$. Although it is relatively straightforward to identify the effect of treatment on surrogates within the experimental sample, identifying the treatment effect on long-term outcome Y is more challenging as the

experimental sample lacks observations of these long-term outcomes, which are only present in the historical sample.

To tackle the challenge of inferring causality from observational data, which may be confounded, we combine the use of surrogates with Pearl's backdoor adjustment. We propose two algorithms that construct a valid set of surrogates and backdoor adjustments to identify the long-term treatment impact. Let $D_{WY}^{\mathcal{G}}$ store the directed paths from W and Y in graph $\mathcal{G}$ and $\psi(p)$ represent the second to last node on path p. For a graph $\mathcal{G}$, Algorithm 3 constructs a valid set of surrogates. To construct a valid set of surrogates, the algorithm examines each directed path from the treatment variable W to the outcome Y within the graph $\mathcal{G}$. For every directed path $p \in D_{WY}^{\mathcal{G}}$, the algorithm selects the second-to-last variable on this path and includes it in the surrogate set $\mathbf{S}_{\mathcal{G}}$. It's important to note that while our algorithm identifies a specific set of surrogates, alternative surrogate sets could also comply with the conditions outlined in Definition 4. Our strategy for surrogate selection is specifically designed to leverage backdoor control variables that adjust for confounding between the surrogates and the outcome, enabling us to represent the long-term treatment effect through a closed-form formula we introduce later in this section.

Using Algorithm 4, we construct an adjustment set that effectively blocks backdoor paths to identify the effects of the surrogates in $\mathbf{S}_{\mathcal{G}}$ on the long-term outcome Y using observational data. Let $B_{SY}^{\mathcal{G}}$ store the backdoor paths relative to $S \in \mathbf{S}_{\mathcal{G}}$ and Y in $\mathcal{G}$. Algorithm 4 systematically examines each surrogate variable $S \in \mathbf{S}_{\mathcal{G}}$ and all backdoor paths in $B_{SY}^{\mathcal{G}}$. For each backdoor path $p' \in B_{SY}^{\mathcal{G}}$, the algorithm identifies the second-to-last variable on this path and includes it into the backdoor adjustment set $\mathbf{Z}_{\mathcal{G}}$ if $noncolliders(p')$ do not intersect with the surrogate set $S \in \mathbf{S}_{\mathcal{G}}$. Proposition 2 below proves that the set $\mathbf{Z}_{\mathcal{G}}$ identified by Algorithm 4 satisfies the backdoor criterion, provided in Definition 3, relative to $(\mathbf{S}_{\mathcal{G}}, Y)$ in graph $\mathcal{G}$, thereby controls for observed confounders when estimating the causal effect of the surrogates $\mathbf{S}_{\mathcal{G}}$ on the outcome $\mathbf{Y}$.

Algorithm 3: Surrogate Selection

Input: $\mathcal{G}, D_{WY}^{\mathcal{G}}$.
Output: $\mathbf{S}_{\mathcal{G}}$.
Initialization: $\mathbf{S}_{\mathcal{G}} = \emptyset$.
1. **for** $p \in D_{WY}^{\mathcal{G}}$:
    $\mathbf{S}_{\mathcal{G}} \leftarrow \mathbf{S}_{\mathcal{G}} \cup \psi(p)$.
2. Return $\mathbf{S}_{\mathcal{G}}$.

PROPOSITION 2. *Let $\mathcal{G}$ be a graph that satisfies Assumptions 3, 4, and 5. Let $\mathbf{S}_{\mathcal{G}}$ and $\mathbf{Z}_{\mathcal{G}}$ be the sets of variables constructed in Algorithms 3 and 4 over the graph $\mathcal{G}$, respectively. Then $\mathbf{Z}_{\mathcal{G}}$ satisfies the backdoor criterion provided in Definition 3 relative to $(\mathbf{S}_{\mathcal{G}}, Y)$ in the graph $\mathcal{G}$.*

Algorithm 4: Backdoor Adjustment Selection

Input: $\mathcal{G}, \mathbf{S}_{\mathcal{G}}, B_{SY}^{\mathcal{G}}$ for all $S \in \mathbf{S}_{\mathcal{G}}$.
Output: $\mathbf{Z}_{\mathcal{G}}$.
Initialization: $\mathbf{Z}_{\mathcal{G}} = \emptyset$.
1. **for** $S \in \mathbf{S}_{\mathcal{G}}$:
  for $p' \in B_{SY}^{\mathcal{G}}$:
    if $noncolliders(p') \cap (\mathbf{S}_{\mathcal{G}} \setminus S) = \emptyset$:
      $\mathbf{Z}_{\mathcal{G}} \leftarrow \mathbf{Z}_{\mathcal{G}} \cup \psi(p')$.
2. Return $\mathbf{Z}_{\mathcal{G}}$.

Let $\mathbf{S}_{\mathcal{G}}$ and $\mathbf{Z}_{\mathcal{G}}$ be the surrogate and backdoor adjustment sets, constructed by Algorithms 3 and 4 for a graph $\mathcal{G}$. We next establish the long-term treatment effect identification strategy using $\mathbf{S}_{\mathcal{G}}$ and $\mathbf{Z}_{\mathcal{G}}$.

PROPOSITION 3. ***(Surrogate Adjustment).*** *Suppose Assumptions 3, 4, 5, and 6 hold. Then, the causal effect of treatment W on the long-term outcome Y for any graph $\mathcal{G}$ is given by*

$$\begin{aligned} P(Y = y \mid D = E, do(W = w)) = \sum_{\mathbf{z}_{\mathcal{G}} \in \mathcal{Z}_{\mathcal{G}}, \mathbf{s}_{\mathcal{G}} \in \mathcal{S}_{\mathcal{G}}} & P(Y = y \mid D = O, \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}, \mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}}) \\ & \times P(\mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}} \mid D = E, W = w, \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}) \\ & \times P(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}} \mid D = E), \end{aligned} \tag{5}$$

where the surrogates $\mathbf{S}_{\mathcal{G}}$ and the backdoor adjustment set $\mathbf{Z}_{\mathcal{G}}$ are obtained via Algorithms 3 and 4 for the graph $\mathcal{G}$. Furthermore, $\mathcal{S}_{\mathcal{G}}$ and $\mathcal{Z}_{\mathcal{G}}$ are the supports of $\mathbf{S}_{\mathcal{G}}$ and $\mathbf{Z}_{\mathcal{G}}$, respectively.

The proposition states that, under the established assumptions, we can identify the causal effect of treatment W on the long-term outcome Y in a closed form using a two-sample framework consisting of both experimental and observational data. Despite Y being unobserved in the experimental sample, the method reformulates the interventional distribution $P(Y = y \mid D = E, \mathrm{do}(W = w))$ using observational distributions that can be estimated directly from either the experimental or observational samples. This proposition also shows that the proposed algorithms for selecting surrogates and backdoor covariates provide a complete identification strategy for all graphs suitable for the surrogacy setting, unlike the backdoor criterion, which may fail to identify causal effects in some graphs (Pearl 2000).

Let $\tau_{\mathcal{G}}$ represent the average treatment effect over graph $\mathcal{G}$, where $\tau_{\mathcal{G}} = \sum_{y \in \mathcal{Y}} y \times \big(P(Y = y \mid D = E, \mathrm{do}(W = 1)) - P(Y = y \mid D = E, \mathrm{do}(W = 0))\big)$. If we had access to the true underlying graph $\mathcal{G}^*$, we could use Proposition 3 to compute the true average treatment effect. However, since the true underlying graph $\mathcal{G}^*$ is unknown, we rely on the graphs returned by the COMB-PC algorithm. Our next result demonstrates that we can recover the true average treatment effect $\tau_{\mathcal{G}^*}$ using the graphs generated by the COMB-PC algorithm.

Theorem 3 ***(Average Treatment Effect Consistency).*** *Suppose Assumptions 1 –7 hold. Let* $\mathbf{G}$ *consist of the graphs returned by the COMB-PC algorithm. Then, we have* $\tau_{\mathcal{G}^*} \in \{\tau_{\mathcal{G}} \mid \mathcal{G} \in \mathbf{G}\}$*, where* $\tau_{\mathcal{G}^*}$ *is the average treatment effect over true underlying graph* $\mathcal{G}^*$*.*

Theorem 3 establishes the accuracy of the proposed causal discovery and treatment effect identification frameworks under the appropriate assumptions. Note that Assumption (7) plays an important role here by ensuring that the graphs returned by the COMB-PC algorithm are within the Markov-equivalence class of the true graph. The next section demonstrates the accuracy of both the discovered graphs and the long-term treatment effect estimates without the oracle assumption, using synthetically generated data.

6. Numerical Experiments with Synthetic Data

In this section, we evaluate the performance of COMB-PC algorithm and the proposed treatment effect identification strategy using synthetic data. We consider four distinct scenarios varying the edge inclusion probability p. The edge inclusion probability p is the fixed probability with which each potential edge between any pair of nodes in a graph is included, thereby determining the overall density of the graph. We use the following probabilities in our study: 0.2, 0.3, 0.4, and 0.5. In each of these scenarios, we generate 200 DAGs, with the number of nodes varying from 10 to 15. We ensure that the generated DAGs satisfy the Assumptions 4 and 5. These DAGs are parameterized as linear Gaussian models and we simulate experimental and observational samples with 50,000 observations. Then, we conduct conditional independence tests using partial correlations, applying a significance threshold of alpha = 0.05, adjusted with a Bonferroni correction to account for multiple hypothesis testing. We use the *pcalg package* (Markus Kalisch et al. 2012) for both generation of DAGs and samples as well as for conducting conditional independence tests.

Table 1 assesses the performance of the COMB-PC algorithm at various edge inclusion probabilities. The MEC column shows the average number of graphs within the Markov equivalence class for each density scenario. As the edge inclusion probability increases from 0.2 to 0.5, the number of graphs within the same Markov equivalence class decreases. As the edge inclusion probability increases, paralleling the decrease in the number of graphs within the Markov equivalence class, we also observe a corresponding reduction in the number of undirected edges. Next, we calculate the average true positive rate (TPR), false positive rate (FPR), true negative rate (TNR), and false negative rate (FNR) for the graphs returned by the COMB-PC algorithm. The true positive rate (TPR), which indicates the algorithm's accuracy in correctly identifying true causal directions, shows a decreasing pattern, dropping from 0.858 at a density of 0.2 to 0.601 at a density of 0.5. This decline suggests a reduced accuracy in identifying causal relationships in denser graphs. Conversely, the false positive rate (FPR) increases with graph density, rising from 0.005 to 0.033, which

points to a higher rate of incorrectly identified causal directions in denser networks. Complementing these trends, the true negative rate (TNR) slightly decreases, and the false negative rate (FNR) significantly increases with higher densities.

Table 1 Evaluating COMB-PC Algorithm's Performance Across Different Edge Densities.

Edge inclusion probability	MEC	Unoriented Edges	Oriented Edges			
			TPR	**FPR**	**TNR**	**FNR**
$p = 0.2$	3.445	1.565	0.858	0.005	0.995	0.142
$p = 0.3$	3.260	1.255	0.813	0.01	0.99	0.187
$p = 0.4$	2.015	0.82	0.786	0.019	0.981	0.214
$p = 0.5$	1.405	0.37	0.601	0.033	0.967	0.399

Having established the accuracy of the COMB-PC algorithm for identifying causal relationships, we next evaluate our proposed treatment effect identification strategy. We begin by calculating the true average treatment effect $\tau_{\mathcal{G}^*}$ for each true underlying graph $\mathcal{G}^*$ by considering all directed paths from the treatment to the outcome and multiplying the true edge coefficients along these paths. We then empirically estimate the treatment effect $\tau_{\mathcal{G}}$ using short-term experimental and long-term observational data. To do this, we employ Algorithms 3 and 4 to identify relevant surrogate and backdoor adjustment sets over the graphs returned by the COMB-PC algorithm. Table 2 summarizes the estimation errors, expressed as $(\tau - \tau^*)/\tau^*$, assessing the deviation of the estimated effects from the true effects across 200 simulations for varying graph densities. The results reveal that estimation errors are minimal in sparser graphs, with a mean error of 0.048 at $p = 0.2$. As graph density increases, mean errors rise, reaching 0.205 at $p = 0.5$. This trend underscores the crucial role of precise causal graph identification, as the diminishing accuracy in causal structure learning within denser graphs significantly compromises the reliability of treatment effect estimations.

Table 2 Average Treatment Effect Estimation Errors Across Varying Graph Densities.

	min	25%	mean	50%	75%	max
$p = 0.2$	-0.186	-0.009	0.048	0.001	0.011	1.113
$p = 0.3$	-0.524	-0.004	0.125	0.006	0.113	1.0
$p = 0.4$	-0.166	-0.005	0.126	0.002	0.115	1.065
$p = 0.5$	-0.233	-0.009	0.205	0.020	0.321	1.021

7. Case Study: Healthy Grocery Shopping

This section presents a case study demonstrating the applicability of our proposed framework for predicting the long-term effects of public policies. It also highlights empirical findings that showcase

how short-term healthy product subsidies can effectively enhance long-term healthy food consumption for specific subsets of users. Many policies employ short-term incentives to shape long-term behaviors. It is crucial to assess whether the costs of these programs are offset by their long-term benefits. We illustrate how our framework enables policymakers to predict these effects in advance and adjust strategies if they are found not to achieve the intended outcomes. We apply our framework to the U.S. Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), focusing on the program's reform that introduced vouchers for healthier food options. Then, we also demonstrate that targeting specific population segments with healthy product subsidies, particularly those with historically low consumption, can lead to lasting positive outcomes. Notably, younger consumers in urban areas are more likely to sustain healthier eating habits even after the subsidies expire.

The remainder of this section is organized as follows. Section §7.1 provides an overview of the data used in the study and describes the experimental setting. Section §7.2 provides evidence regarding the persistent effects of the subsidies using raw data, while Section §7.3 estimates the long-term treatment effect using "future data", in which the long-term outcome is observed, through the difference-in-difference method. In Section §7.4, we strategically overlook the "future data" and predict the long-term treatment effect using the proposed strategies combining a short-term sample with a historical sample. Finally, Section §7.5 provides our novel empirical findings.

7.1. Data Description and Experimental Setup

The WIC is a federal assistance program in the United States that provides nutrition and health support to low-income pregnant women, new mothers, and young children up to age of five. In 2016, about 8 million people participated in WIC each month, which made up 6% of all spending on food and nutrition assistance in the US. WIC plays an important role in supporting the nutritional needs of low-income families and helping to reduce the risk of low birth weight, promote child growth, and encourage healthy eating habits. In 2009, the WIC program made big changes to the foods it provides. The goal was to make the WIC food packages match the latest dietary recommendations. WIC was originally established to help low-income families avoid malnutrition, but some people were worried that it might be contributing to childhood obesity. The 2009 reform added new food options, like whole-grain products, fruits, and vegetables (e.g., whole-wheat bread). Among these, 100% whole-wheat bread was a significant addition, with post-reform vouchers restricted to this product.

Specifically, in this study, we focus on the consumption of 100% whole-wheat bread, as it was one of the most targeted product categories by the reform (Hinnosaar 2023). To this end, we analyze the impact of the reform on the total quantity of whole-wheat bread purchases aggregating across

brands, package sizes, and other product characteristics. We use data from the NielsenIQ Consumer Panel[7] to track consumer purchases. The data is representative of the U.S. population and it covers 14 years, from 2006 to 2019. The households in the study were asked to scan all the groceries they bought for personal consumption at home. The data collected includes the UPC code, quantity, and price of each item purchased, as well as demographic information about the households. Most importantly for our purposes, the dataset has provided yearly information on households' self-reported WIC status since 2006. In preparing the data for this study, we followed a preprocessing approach similar to that described in the paper by Hinnosaar (2023), to which we direct the reader for detailed information on the data and the descriptive statistics.

7.1.1. Experimental Setup. In our experimental setup, we assume the ability to only access customer consumption data for the first year post-treatment (i.e., we observe customers purchases within the first year after they stops receiving subsidies for healthy products), aiming to predict their consumption in the following second and third years post-treatment. To this end, we intentionally ignore the data from the second and third years post-treatment, focusing instead on the short-term impact of this reform observed in the first year post-treatment to estimate the effects of these subsidies on healthy product consumption in later years. Recall that our framework relies on leveraging both short-term experimental data and observational data when making long-term predictions of the effect of the reform (i.e., treatment intervention). The crucial aspect of our long-term treatment prediction framework is identifying short-term surrogate variables that mediate the treatment effect on long-term healthy product consumption and are observed in both short-term experimental data and long-term historical data. With that in mind, we next explain the main logic behind identifying candidates for these surrogate variables and how they can be used to link short-term experimental data with long-term historical data based on our particular context of subsidizing healthy product consumption.

The primary effect of the WIC reform on the consumption of healthy bread is driven by the monetary stimulus, which essentially means that customers participating in the WIC program receive price discounts when purchasing healthy food. Hence, we believe that the impact of subsidies from the WIC reform for healthy products manifests through increased discount-seeking behavior after the consumers stop receiving the WIC stimulus. Therefore, variables such as the frequency of healthy bread purchases with deals and the frequency of healthy bread purchases with coupons can be considered as 'potential' surrogates that are observable in the historical data. Furthermore, we believe this setting is suitable for the surrogacy framework, as the monetary stimulus from the WIC reform

[7] The dataset is sourced from Nielsen Consumer LLC and marketing databases accessible through the NielsenIQ Datasets at the Kilts Center for Marketing Data at The University of Chicago Booth School of Business https://www.chicagobooth.edu/research/kilts/datasets/nielsenIQ-nielsen.

is likely to affect long-term consumption of healthy bread by first impacting short-term consumption patterns. This change is expected to occur in the quarters immediately following the termination of the intervention. As a result, variables such as the proportion of healthy bread purchases made with or without discounts in the quarters after the intervention could serve as potential mediators in our analyses.

Table 3 presents the short-term variables we used in our empirical analysis. In the third column of this table, we specify the set of variables that are observed one year after the treatment intervention ends. Most of those variables are aggregated at a customer-level: (a) *Fraction healthy bread* Q_i is the fraction of healthy bread in the overall bread purchases over a quarter i of the first year after the subsidizing period ends; (b) *Average number healthy bread deals* Q_i is the number of times when the purchased healthy bread was under promotion over a quarter i of the first year after the subsidizing period ends; (c) *Average value healthy bread coupon* Q_i is the average value of the coupon that was applied to healthy bread purchased over a quarter i of the first year after the subsidizing period ends; (d) *Healthy bread price* Q_i is the average price of a healthy bread purchased over a quarter i of the first year after the subsidizing period ends; (e) *Fraction healthy bread* is the fraction of healthy bread in the overall bread purchases during the subsidizing period; (f) *Average number healthy bread deals* is the number of times when the purchased healthy bread was under promotion during the subsidizing period; (g) *Average value healthy bread coupon* is the average value of the coupon that was applied to healthy bread purchased during the subsidizing period; and (h) *Healthy bread price* is the average price of a healthy bread purchased during the subsidizing period. Then, we also have the time-varying household characteristics such as the logarithm of income (i.e., *Log household income*), household size, age, education, and indicator variable if less than an 18-year-old kid is part of a household (i.e., *Children below 18*) both for the historical and experimental datasets. Finally, in the last column of Table 3, we have the long-term outcome variable *Fraction healthy bread long term* which we want to predict since it is not observed in the short term. More specifically, *Fraction healthy bread long term* is the fraction of healthy bread purchases in the second and third years after the subsidy ends.

To complement the short-term experimental data, we use three years of historical data (2006, 2007, and 2008) on grocery purchases where we have access to the information on products' promotions, coupons, price, indicator variable if a purchased product is a healthy (or non-healthy) bread as well as to the time-varying household characteristics such as the income, household size, age, education, and indicator variable if less than an 18-year-old kid is part of a household. Overall, the integration of experimental and historical datasets allows us to assess the predictive capacity of our framework. We then test the predictive capacity of our model by comparing its forecasts

against the actual consumption data from the second and third years post-treatment—after customers finished receiving subsidies for healthy products.[8] Thus, we use this experimental setting to empirically validate our framework's ability to accurately forecast the long-term effects of healthy product subsidies, leveraging short-term observations as surrogate variables.

Table 3 Description of Our Setup in the Context of the WIC Reform.

Sample	Treatment (Observed)	Variables within First Year after Subsidy Ends (Observed)	Long-term Outcome (Unobserved)
Experimental Sample	WIC vouchers after the 2009 reform	Average number healthy bread deals, Average number healthy bread deals Q1-Q4, Average value healthy bread coupon, Average value healthy bread coupon Q1-Q4, Fraction healthy bread, Fraction healthy bread Q1-Q4, Healthy bread price, Healthy bread price Q1-Q4, Household size, Age, Log household income, Children below 18, Education	Fraction healthy bread long term

7.2. Model Free Evidence

In this subsection, we present key findings obtained from the raw data (i.e., model-free evidence), shedding light on the plausible direction of the effect size and setting the stage for a more rigorous examination through econometric modeling in the next subsection. As Hinnosaar (2023) has already demonstrated that this reform does not affect the long-term consumption of healthy bread for the entire population, we focus on consumers with historically low healthy bread purchases.[9] Specifically, we focus our analysis on customers within the treatment group who are in the bottom half of households based on their share of healthy bread in total bread purchases during the pretreatment period.[10] To this end, Figure 5 presents model-free evidence on persistent effects of these subsidies when we focus only on customers with historically low healthy bread purchases. The figure illustrates variations in the percentage of healthy bread within the overall bread purchases across distinct

[8] Note that our goal is to evaluate the long-term effect of the treatment intervention, which is quantified by the consumption of the healthy bread during the second and third years post-treatment when only having access to the consumption of the healthy bread during the first year post-treatment. In other words, we evaluate the proposed framework's ability to estimate the long-term effect of the treatment soon after the experiment concludes, without waiting for many years to estimate this long-term effect.

[9] We were also able to confirm in our empirical setup that healthy product subsidies do not have a long-lasting impact on the entire population. In the interest of space, we omit this analysis and instead refer to the findings of Hinnosaar (2023).

[10] As the goal of this policy is to improve the consumption level of healthy bread, we focus our analysis on customers with the lowest proportion of healthy bread purchases during the pretreatment period. This group has the least healthy consumption habits, making it fair and logical to prioritize them for policy intervention testing, especially since the policy did not achieve its long-term objective when applied to all WIC households according to Hinnosaar (2023).

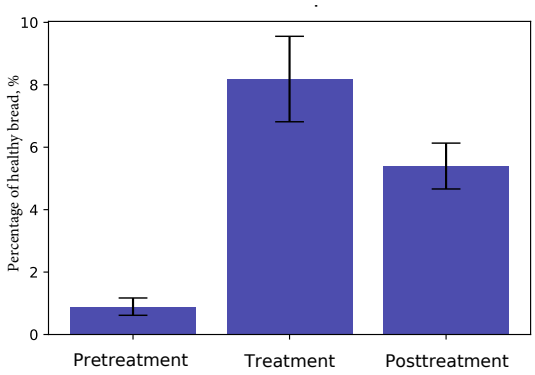

Figure 5 Model free evidence. In this figure, we present the percentage of healthy bread for the treated customers before the treatment, during the treatment, and after the treatment when we only focus on the bottom 50% of treated customers who had a relatively low percentage of healthy bread in their purchases during the pretreatment period of time. The brackets represent 99% confidence interval.

periods: the pretreatment period (prior to receiving WIC vouchers for healthy bread purchases), the treatment period (during the receipt of WIC benefits), and the posttreatment period (any time after finishing the receipt of WIC vouchers).

Since the households in this sample were selected based on their low consumption of healthy bread, the average percentage of healthy bread in their purchases is 0.9%, which is much lower than the average percentage of healthy bread consumed by all households in our study during the pretreatment period of time (i.e., 10.7%). What is even more interesting is that, for this specific set of households, the proportion of healthy bread in overall bread purchases in the long term (after finishing the receipt of WIC vouchers) remains notably higher than the pretreatment level (5.4% versus 0.9%), presenting a contribution to the results obtained by Hinnosaar (2023) in the same setting where the authors show that this reform does not change the long-term consumption of the healthy bread for the overall population. Therefore, this model-free evidence suggests that WIC reform might be actually effective in changing the household persistence to healthy bread when focusing on select households who had relatively low consumption of healthy bread.

7.3. Estimating the Long-Term Effect of WIC Reform with "Future Data"

The goal of this section is to analyze the long-term effects of healthy food subsidies on WIC households using an experimental sample that actually includes the long-term outcome, referred to as "future data". As discussed previously, the analysis relies on the 2009 WIC program reform with a focus on the whole wheat bread consumption. It compares changes in purchases associated with the start and end of WIC voucher receipt in two distinct household groups: those receiving vouchers from the old program (control households) and those from the new program (treatment households). While these two groups of households share similarities, the vouchers they receive differ. First, we

evaluate the long-term impact of WIC policy reform by employing the following regression model that leverages the empirical methodology effectively utilized by Hinnosaar (2023).

$$\begin{aligned} Y_{it} = {} & \beta_1 WIC_{it} + \beta_2 AfterWICYears1_{it} + \beta_3 AfterWICYears2,3_{it} + \beta_4 ReformedWIC_{it} \\ & + \beta_5 AfterReformedWICYears1_{it} + \beta_6 AfterReformedWICYears2,3_{it} + \beta_7 AfterWICYear4^{+}_{it} \\ & + \beta_8 AfterReformedWICYear4^{+}_{it} + X_{it}\eta + \delta_i + \gamma_t + \varepsilon_{it}, \end{aligned} \tag{6}$$

where Y_{it} is the percentage of healthy bread in overall bread purchases of household i in time period t, WIC_{it} is a binary variable and it is equal to 1 if a household i receives WIC vouchers in period t, $AfterWICYears1_{it}$ indicates that household i finished using WIC vouchers when making the purchases up to one year earlier, $AfterWICYears2,3_{it}$ indicates that household i finished using WIC vouchers when making the purchases two to three years earlier, $ReformedWIC_{it}$ is an indicator variable and it is equal to 1 if and only if $WIC_{it} = 1$ and t corresponds to the time period after the reform, $AfterReformedWICYears1_{it}$ indicates that household i finished using reformed WIC vouchers when making the purchases up to one year earlier, $AfterReformedWICYears2,3_{it}$ is the estimate of the long-term effect size. Then, X_{it} is included into the aforementioned regression specification to capture time-varying household characteristics such as the logarithm of income, household size, age, education, and indicator variable if less than a 18 years old kid is part of a household. We also include δ_i and γ_t to capture household and time period fixed effects. Finally, dummy variables $AfterWICYear4^{+}_{it}$ and $AfterReformedWICYear4^{+}$ indicate whether households received WIC or reformed WIC vouchers four or more years earlier. As it is standard in the literature to increase statistical power in this way, we do not drop the observations three or more years after receiving the vouchers but we do not report the coefficient estimates for these dummy variables because the balanced panel only includes data up to three years after receiving vouchers (Hinnosaar 2023).

Table 4 displays the results obtained by estimating regression specification (6) based on the bottom half of households that initially (i.e., during the pretreatment period) purchased the least amount of healthy bread. The table shows the treatment (i.e., 2009 WIC reform) leads to a 2.63 percentage point increase in the long-term proportion of healthy bread purchases.

7.4. Treatment Effect Prediction using Surrogates without "Future Data"

In this section, we forecast the long-term effect using data from the first year only, as described in §7.1.1, and compare them with the long-term effect estimated in the previous subsection. Following our framework, the first step is to obtain the causal graph by using the COMB-PC algorithm that integrates experimental and observational samples of the data to find the underlying causal

Model Dependent Variable	(1) Bread healthy %
Reformed WIC	6.6006***
	(1.0893)
After reformed WIC year 1	6.9820***
	(1.4287)
After reformed WIC years 2 and 3	2.6359**
	(1.1919)
WIC	1.0267
	(0.7454)
After WIC year 1	1.6887*
	(1.0211)
After WIC years 2 and 3	2.8713***
	(1.0758)
Controls	Yes
Year-quarter fixed effects	Yes
Household fixed effects	Yes
R^2	0.28724
Observations	5,921

Table 4 **Main analyses. We estimate the long-term effect of treatment (based on years 2 and 3) when we only focus on the bottom 50% of treated customers who had a relatively low percentage of healthy bread in their purchases during the pretreatment period of time.**

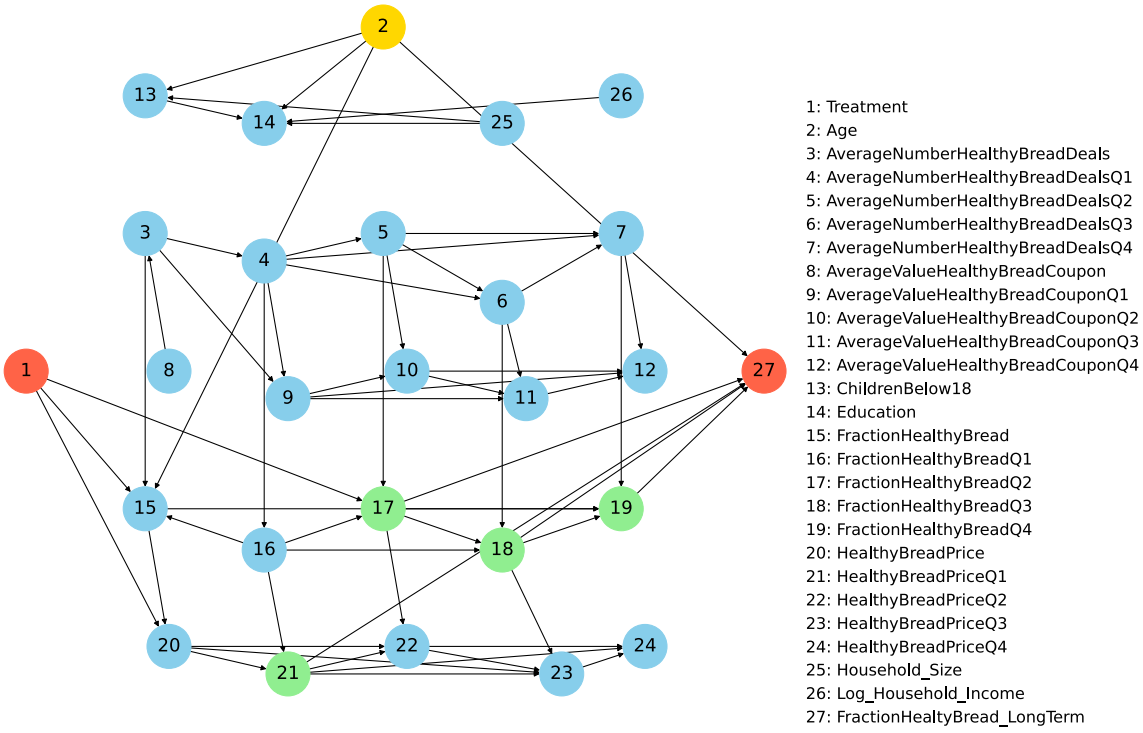

Figure 6 **Causal graph generated using the COMB-PC algorithm.**

relationships. Figure 6 illustrates the graph generated by the COMB-PC algorithm where we have the set of the variables provided in Table 3. Interestingly, although the COMB-PC algorithm usually produces a Markov equivalence class, in our case, we obtained a unique graph. Having this graph, we first apply Algorithm 3 to identify valid surrogate variables. *Fraction healthy bread Q2*, *Fraction healthy bread Q3*, *Fraction healthy bread Q4*, and *Healthy bread price Q1* variables are identified as valid surrogates because every directed path from the *Treatment* to the long-term outcome variable (i.e., *Fraction healthy bread long term*) intersects with at least one of these variables. Next, Algorithm 4 is used to identify the confounding variables between the aforementioned surrogates and the long-term outcome variable (i.e., *Fraction healthy bread long term*). The outcome of Algorithm 4 indicates that controlling for the *Age* variable is sufficient to address confounding issues.

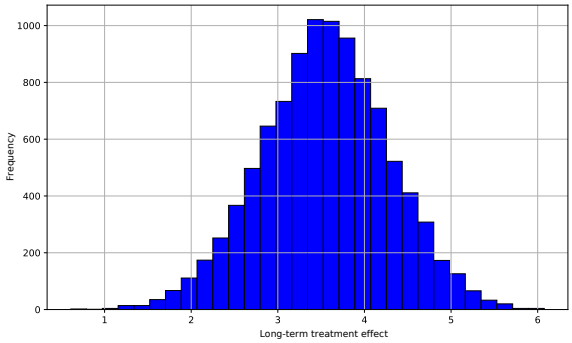

Figure 7 **The long-term treatment effect estimation results based on 10,000 bootstrapped samples using surrogate adjustment.**

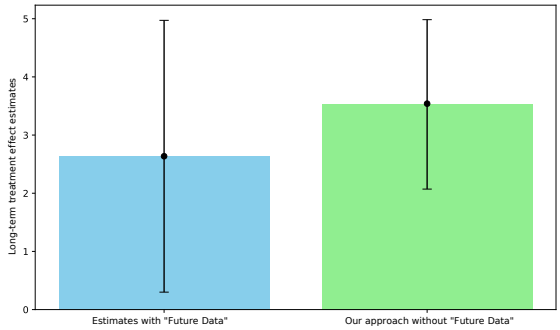

Figure 8 **Comparison of long-term effect estimations using the standard empirical method based on "future data" (see the left panel) versus our surrogate variable framework without using "future data" (see the right panel). The bars represent 95% confidence intervals.**

After building the causal graph and identifying surrogate variables as well as backdoor variables, we estimate the long-term treatment effect by invoking Proposition (3) over 10,000 bootstraped samples. Figure 7 displays the distribution of long-term treatment effect estimates derived from the

bootstrapped samples. Furthermore, the right panel of Figure 8 elaborates on our result, indicating that the treatment is forecasted to result in a 3.55 percentage point increase in the proportion of healthy bread consumed by customers in our data sample over the long term, with a 95% confidence interval ranging from 2.07% to 4.98%. Then, the left panel illustrates the results obtained from the preceding section, where we assume access to the "future data" and estimate that the treatment leads to a 2.63 percentage point increase in the proportion of healthy bread consumed by customers in our data sample over the long term, with a 95% confidence interval ranging between 0.3% and 4.9%. The overlapping confidence intervals suggest that one can effectively predict the long-term treatment effect by leveraging our framework with surrogate variables.

7.4.1. Discussion and Policy Implications. In the previous section, we demonstrated that our proposed method for identifying surrogate variables using the causal discovery framework performs exceptionally well in practice. By introducing a *novel* nonparametric end-to-end framework, we not only uncover surrogates in a data-driven manner, but also leverage these surrogates to estimate the long-term impact of treatment interventions by combining short-term experimental data with long-term historical data. We believe our method offers valuable insights for decision-makers by uncovering relationships that may not be readily apparent or that exceed conventional intuition and expert judgment, thereby facilitating a rigorous, data-driven approach to identifying surrogate variables that mediate treatment effects. Additionally, this methodology supports proactive policy adjustments or terminations, thereby minimizing potential significant costs if the anticipated long-term goals are not achieved. Ultimately, this enables policymakers to efficiently assess both immediate results and forecasted long-term outcomes, facilitating a more streamlined decision-making process without significant delays.

7.5. Empirical Findings: Heterogenous Treatment Effect

Our empirical findings so far indicate that while subsidies on healthy products do not significantly affect long-term consumption of healthy bread overall, this policy can be particularly effective for a specific group of users who have had limited prior exposure to healthy products. This insight offers critical implications for policymakers seeking to encourage healthier eating habits among consumers. In this subsection, we delve further into the analysis, exploring the key characteristics that may influence the long-term effectiveness of such subsidies, aiming to better understand which factors drive the policy's success for certain groups over others. This deeper exploration follows a similar spirit to the analysis in Section § 7.3 and can help refine targeted interventions and optimize health-focused subsidy programs for lasting impact. To this end, we interact the treatment effect with two customer-specific attributes: Age[11], representing the average age of household heads, and

[11] NielsenIQ dataset categorizes the age of the household head as an integer ranging from 1 to 9. These values correspond to the following age groups: 1 represents under 25, 2 corresponds to 25-29, 3 to 30-34, 4 to 35-39, 5 to 40-44, 6 to 45-49, 7 to 50-54, 8 to 55-64, and 9 to 65+.

Density, which represents the population density of the household's area of a residence based on zipcode-level[12] data. More specifically, we estimate the following regression specification:

$$\begin{aligned} Y_{it} = {} & \beta_1 WIC_{it} + \beta_2 AfterWICYears1_{it} + \beta_3 AfterWICYears2,3_{it} + \beta_4 ReformedWIC_{it} \\ & + \beta_5 AfterReformedWICYears1_{it} + \beta_6 AfterReformedWICYears2,3_{it} + \beta_7 AfterWICYear4^{+}_{it} \\ & + \beta_8 AfterReformedWICYear4^{+}_{it} + \beta_9 Var \times AfterWICYears1_{it} + \beta_{10} Var \times AfterWICYears2,3_{it} \\ & + \beta_{11} Var \times ReformedWIC_{it} + \beta_{12} Var \times AfterReformedWICYears1_{it} \\ & + \beta_{13} Var \times AfterReformedWICYears2,3_{it} + X_{it}\eta + \delta_i + \gamma_t + \varepsilon_{it}, \end{aligned} \tag{7}$$

where *Var* is either *Age* or *Density* and all the other variables have the same definition as in the Equation (6). In what follows below, we examine the role of these two customer-specific factors where *Bread healthy %* is the dependent variable.

7.5.1. Household's Age. First of all, note that it is not clear ex-ante whether healthy product subsidies will have a more lasting effect on older or younger consumers' long-term purchasing habits, particularly once the subsidies are removed. On the one hand, older consumers may be more likely to adopt and sustain healthier purchasing behaviors because they tend to be more focused on long-term health outcomes. With age, concerns over chronic health conditions such as heart disease and diabetes become more pressing, potentially incentivizing older individuals to increase the fraction of healthy products in their diet even after the subsidies end. On the other hand, it is equally plausible that younger consumers could continue buying healthy products after the subsidies end. Younger consumers may develop lasting habits and preferences for healthy foods during the subsidy period. Behavioral economics suggests that repeated exposure to healthier products, especially when combined with positive reinforcement like savings, could lead to habit formation (Cawley and Ruhm 2011, Roberto and Kawachi 2015). This could drive younger consumers to continue purchasing healthy items even after the incentive is gone, as their taste preferences or perceptions of value may shift during the subsidy period. Furthermore, younger consumers may be more adaptable and open to new food choices, especially if the experience with subsidized healthy products is positive. Over time, these changes could result in a sustained shift in their buying behavior. As such, it is difficult to predict in advance which age group is more likely to maintain long-term consumption of healthy products once subsidies are removed. To examine how the impact of the healthy product subsidies depends on the age of a household, we estimate the regression specification (7) where our moderator is *Age*.

Column (1) of Table 5 presents the regression results, where the interaction term, *Age* × *After reformed WIC years 2 and 3*, is particularly insightful. The negative coefficient suggests that the

[12] Zipcode-level density information can be accessed via this link: https://www.fourfront.us/data/datasets/.

younger the household, the more likely the healthy product stimulation policy impacts the long-term consumption of healthy products. This might imply that younger households are more responsive to the policy changes, potentially forming lasting habits due to their adaptability and openness to new purchasing behaviors. These results indicate that younger households might be quicker to integrate healthier products into their regular shopping routines, even after the policy incentives phase out. This could be attributed to habit formation mechanisms (Cawley and Ruhm 2011), where repeated exposure to subsidized healthy products fosters a sustained preference over time. This finding has significant policy implications, as it highlights the importance of considering demographic factors such as age when designing and evaluating the long-term effectiveness of public health interventions like subsidies for healthy products. Younger households may represent a key target for such policies due to their potential for lasting behavioral changes.

7.5.2. Population Density. The next customer-level attribute we examine is the population density of the area of residence of a household. This moderator plays a crucial role in bridging the findings of this paper with the broader food desert literature (Allcott et al. 2019), where a central debate revolves around how the availability and variety of healthy food options influence the nutritional quality of consumers' diets. This exploration is particularly relevant for policymakers aiming to address nutritional disparities by increasing healthy food availability in low-access regions. To this end, column (2) of Table 5 reports the regression results using *Bread healthy %* as the dependent variable and includes *Density* as a moderator. The coefficient for the interaction term, *Density* $\times$ *After reformed WIC years 2 and 3*, indicates that households residing in areas with higher population density experience a greater long-term impact from healthy product subsidies on their consumption of healthy bread. This suggests that households in densely populated areas may benefit more from these policies. One potential explanation for this finding is that areas with higher population density often offer a wider variety of product choices, including healthier options. With greater accessibility and convenience in purchasing healthy products, consumers in these areas may find it easier to incorporate them into their regular diets. This increased exposure and convenience could contribute to stronger habit formation, encouraging sustained consumption of healthy products even after the healthy product subsidies are removed. Additionally, the competitive market in densely populated regions might drive retailers to promote healthier choices more actively, further reinforcing this behavior change among consumers.

7.6. Discussion of the Empirical Findings and Policy Implications

Promoting healthy product choices has become increasingly imperative, with government agencies recognizing its pivotal role in public health (An 2013, Allcott et al. 2019). To address the rising concerns surrounding diet-related health problems, government initiatives have been implemented to

Model	(1)	(2)
Dependent Variable	Bread healthy %	Bread healthy %
Age × *Reformed WIC*	-1.1199**	
	(0.4879)	
Age × *After reformed WIC year 1*	-1.3576*	
	(0.6982)	
Age × *After reformed WIC years 2 and 3*	-1.0944**	
	(0.4514)	
Age × *WIC*	0.474	
	(0.3039)	
Age × *After WIC year 1*	-0.0471	
	(0.5022)	
Age × *After WIC years 2 and 3*	0.3726	
	(0.3926)	
Density × *Reformed WIC*		0.0010***
		(0.0002)
Density × *After reformed WIC year 1*		0.0036***
		(0.0007)
Density × *After reformed WIC years 2 and 3*		0.0028***
		(0.0008)
Density × *WIC*		-0.0001
		(0.0001)
Density × *After WIC year 1*		-0.0032***
		(0.0007)
Density × *After WIC years 2 and 3*		-0.0029***
		(0.0008)
Reformed WIC	12.3731***	4.7056***
	(2.9546)	(1.0820)
After reformed WIC year 1	14.2257***	3.3932**
	(4.6471)	(1.5891)
After reformed WIC years 2 and 3	8.7881***	-0.2437
	(2.9645)	(1.3206)
WIC	-1.4266	1.9488***
	(1.7958)	(0.7370)
After WIC year 1	2.0346	4.8348***
	(3.3045)	(1.1900)
After WIC years 2 and 3	0.6649	5.8439***
	(2.5276)	(1.1921)
Controls	Yes	Yes
Year-quarter fixed effects	Yes	Yes
Household fixed effects	Yes	Yes
R^2	0.28972	0.29684
Observations	5,921	5,921

Table 5 **Heterogeneous treatment effect analysis. The moderators are the average age of household heads in the first column and the population density of the household's area in the second column.**

encourage healthier choices through both supply-side and demand-side subsidies. Many researchers in the public health sector, policymakers, and advocates assert that the existence of food deserts constitutes a crucial factor contributing to unhealthy eating habits (Shannon 2014). This has led to governments at both the federal and local levels investing millions of dollars annually in supply-side policies that provide financial support and assistance to grocery stores operating in underserved

areas. However, some researchers argue that supply-side strategies, such as promoting the opening of grocery stores in underserved areas, may not have a significant impact on people's eating habits (Allcott et al. 2019). Instead, they suggest an alternative approach: providing subsidies for healthy foods (e.g., WIC reform). These initiatives aim to change consumer behavior by making healthy food choices more readily available and less expensive. By providing subsidies for items like whole wheat bread, governments hope to encourage people to adopt healthier eating habits and reduce the detrimental effects of unhealthy food consumption on public health. However, as it was mentioned above, Hinnosaar (2023) shows that even the effect of the demand-side policies (i.e., healthy product subsidies) is mixed in the long run. Despite extensive evidence in the literature demonstrating the persistence and challenges of altering nutritional choices (Ma et al. 2013, Atkin 2013, Bryan et al. 2016, Biesbroek et al. 2023), our empirical contribution to this research field reveals that healthy product subsidies can have a more enduring impact on individuals with historically low health product purchases, even though these subsidies may not be effective for the entire population as it was shown in many existing papers.

Given the considerable efforts invested by government agencies to identify the most effective methods for encouraging healthy eating behaviors among consumers, the findings presented in this paper can provide valuable insights into which customer segments would benefit most from healthy product subsidies, ensuring a more enduring and sustainable impact. Therefore, our primary policy suggestion is to restrict subsidies to individuals with limited prior exposure to healthy products. This group is an ideal target for this policy due to the following two potential mechanisms. First, the subsidies for healthy products could have an *informational effect*. By receiving these subsidies, these customers may gain knowledge about healthy eating and develop a more positive perception of the social status associated with making healthier food choices. Second, there is an increased likelihood that healthy products were never included in the individuals' consideration sets, preventing them from evaluating such products (i.e., *consideration set expansion effect*). Consequently, after trying them for the first time with the financial incentives offered by the government program, these customers might discover a preference for these products and subsequently continue to purchase them consistently in the future. Similarly, we posit that customers who have historically made substantial purchases of healthy products may only exhibit a short-term response to subsidies for such products. This is because they are likely already cognizant of the benefits of healthy eating and have optimized the proportion of healthy products to achieve a balance between health outcomes and the variety of products they consume. Hence, as shown by Hinnosaar (2023), these customers increase their consumption during the subsidy period but subsequently decrease their consumption post-reform, keeping the same average long-term fraction of healthy bread in overall bread purchases. The latter observation is unsurprising, as it mirrors consumer reactions to price discount promotions.

It is a well-established phenomenon that consumers tend to amplify brand purchases in the short term during promotional periods, but the long-term effects are often either nonsignificant or negative (Mela et al. 1997).

Our analysis reveals that the effectiveness of the food policy is influenced by both the age of consumers and the population density of their residential areas. Specifically, younger consumers tend to sustain healthier eating habits in the long term, even after the healthy product subsidies are discontinued. This finding implies that while the current policy effectively fosters lasting behavior changes among younger populations, older consumers may require a different set of interventions, such as alternative incentives or behavioral nudges, to encourage similar long-term dietary improvements. These insights underscore the importance of tailoring public health strategies to different demographic segments, as distinct consumer groups respond differently to short-term policy measures aimed at promoting healthy food consumption. A 'one-size-fits-all' approach may not be as effective across all age groups, and further exploration into targeted interventions for older populations could enhance the overall impact of these policies.

Moreover, our results emphasize that policymakers should account for geographic factors, such as population density, to maximize the policy's reach and efficiency. In this context, our findings suggest that the policy is more likely to achieve its long-term objectives when targeted at consumers residing in areas with higher population density. However, this focus could inadvertently widen the gap in dietary health between urban and rural areas, potentially exacerbating the existing disparity seen in some rural regions, commonly referred to as "food deserts". In the long term, this could lead to greater inequalities in the consumption of healthy foods, with urban areas benefiting more from the policy while rural and lower-density regions continue to lag behind. To address this concern, government policies need to be adapted to ensure that the positive effects of healthy food subsidies reach all communities equally. This may involve creating tailored interventions for rural areas, such as improving access to healthy food options through infrastructure investments, transportation solutions, or alternative subsidy models that address the unique challenges of food deserts. By implementing more region-specific strategies, policymakers can work toward reducing the urban-rural divide and achieving a more equitable impact on public health outcomes across diverse populations.

8. Conclusion, Limitations, and Future Research

In this paper, we develop a framework for estimating long-term treatment effects by integrating short-term experimental data with long-term historical data. Our study makes several contributions to the growing body of literature on surrogacy. First, we present the COMB-PC algorithm, an

innovative causal structure learning algorithm that seamlessly integrates experimental and observational data to reveal the underlying causal relationships among variables. This algorithm, comprising stages of skeleton discovery and edge orientation, adapts traditional causal structure learning methods to our specific framework, highlighting its flexibility and applicability. Building on this, we develop a method to estimate the average long-term treatment effect using surrogate variables, a novel approach that effectively utilizes surrogate variables and backdoor adjustments to bridge the gap between short-term data and long-term outcomes. Our numerical experiments validate this framework, showing its capability to accurately identify causal relationships and estimate treatment effects across various graph densities. Additionally, a real-world case study empirically confirms the framework's effectiveness in forecasting the long-term impact of health product subsidies on consumer behavior, with our findings suggesting that targeted subsidies could promote sustained health-conscious shopping among consumers who historically make fewer healthy grocery purchases. This research not only bridges surrogacy framework with causal discovery but also offers practical insights for designing health-related policy interventions.

There are numerous potential directions for future research. One particular direction is to extend the current framework to incorporate confounding factors and non-random treatment assignment into the estimation framework, to improve causal inference from observational data. Alternatively, the proposed framework can be extended to predict the "long-term" treatment effect of long-term interventions instead of predicting the "long-term" treatment effect of short-term interventions. Another promising direction for future work involves determining the optimal amount of short-term data required to make accurate long-term predictions. In other words, it would be valuable to determine the optimal timing at which short-term data becomes sufficient for making long-term forecasts with our causal structure learning framework.

A promising avenue for future research involves validating the long-term effectiveness of healthy product subsidies among customers with historically low consumption of healthy products through randomized control trials (RCTs). Additionally, it would be valuable to explore further the observed tendency of younger consumers and those in densely populated urban areas to sustain healthier eating habits over time. By verifying these patterns, we can better understand the factors driving long-term behavior change. Future studies could also uncover new insights and nuanced factors that may influence the sustained impact of short-term healthy product subsidies. For instance, different demographic, socioeconomic, or geographic characteristics might play a role in determining how effectively these policies foster long-term healthy habits. Understanding these intricate details will be critical for designing more precise, data-driven interventions aimed at promoting healthy eating across diverse populations.

References

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, Molly Schnell. 2019. Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics* **134**(4) 1793–1844.

An, Ruopeng. 2013. Effectiveness of subsidies in promoting healthy food purchases and consumption: a review of field experiments. *Public health nutrition* **16**(7) 1215–1228.

Anderer, Arielle, Hamsa Bastani, John Silberholz. 2022. Adaptive clinical trial designs with surrogates: When should we bother? *Management science* **68**(3) 1982–2002.

Athey, Susan, Raj Chetty, Guido Imbens. 2020. Combining experimental and observational data to estimate treatment effects on long term outcomes. *arXiv preprint arXiv:2006.09676* .

Athey, Susan, Raj Chetty, Guido W Imbens, Hyunseung Kang. 2019. The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. Tech. rep., National Bureau of Economic Research.

Atkin, David. 2013. Trade, tastes, and nutrition in india. *American economic review* **103**(5) 1629–1663.

Bareinboim, Elias, Judea Pearl. 2012. Causal inference by surrogate experiments: z-identifiability. *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*. 113–120.

Battocchi, Keith, Eleanor Dillon, Maggie Hei, Greg Lewis, Miruna Oprescu, Vasilis Syrgkanis. 2021. Estimating the long-term effects of novel treatments. *Advances in Neural Information Processing Systems* **34** 2925–2935.

Bayati, Mohsen, Andrea Montanari, Amin Saberi. 2018. Generating random networks without short cycles. *Operations Research* **66**(5) 1227–1246.

Biesbroek, Sander, Frans J Kok, Adele R Tufford, Martin W Bloem, Nicole Darmon, Adam Drewnowski, Shenggen Fan, Jessica Fanzo, Line J Gordon, Frank B Hu, et al. 2023. Toward healthy and sustainable diets for the 21st century: Importance of sociocultural and economic considerations. *Proceedings of the National Academy of Sciences* **120**(26) e2219272120.

Bojinov, Iavor, David Simchi-Levi, Jinglong Zhao. 2023. Design and analysis of switchback experiments. *Management Science* **69**(7) 3759–3777.

Bryan, Christopher J, David S Yeager, Cintia P Hinojosa, Aimee Chabot, Holly Bergen, Mari Kawamura, Fred Steubing. 2016. Harnessing adolescent values to motivate healthier eating. *Proceedings of the National Academy of Sciences* **113**(39) 10830–10835.

Caro, Felipe, Jérémie Gallien. 2012. Clearance pricing optimization for a fast-fashion retailer. *Operations research* **60**(6) 1404–1422.

Caro, Felipe, Victor Martínez-de Albéniz. 2012. Product and price competition with satiation effects. *Management Science* **58**(7) 1357–1373.

Cawley, John, Christopher J Ruhm. 2011. The economics of risky health behaviors. *Handbook of health economics*, vol. 2. Elsevier, 95–199.

Chen, Yiwei, Vivek F Farias, Nikolaos Trichakis. 2019. On the efficacy of static prices for revenue management in the face of strategic customers. *Management Science* **65**(12) 5535–5555.

Claassen, Tom, Tom Heskes. 2010. Learning causal network structure from multiple (in) dependence models .

Cooper, Gregory F, Changwon Yoo. 2013. Causal discovery from a mixture of experimental and observational data. *arXiv preprint arXiv:1301.6686* .

Dekimpe, Marnik G, Dominique M Hanssens, Jorge M Silva-Risso. 1998. Long-run effects of price promotions in scanner markets. *Journal of econometrics* **89**(1-2) 269–291.

Eberhardt, Frederick. 2017. Introduction to the foundations of causal discovery. *International Journal of Data Science and Analytics* **3**(2) 81–91.

Eberhardt, Frederick, Nur Kaynar, Auyon Siddiq. 2024. Discovering causal models with optimization: Confounders, cycles, and instrument validity. *Management Science* .

Esenduran, Gökçe, Lauren Xiaoyuan Lu, Jayashankar M Swaminathan. 2020. Buyback pricing of durable goods in dual distribution channels. *Manufacturing & Service Operations Management* **22**(2) 412–428.

Farias, Vivek, Andrew Li, Tianyi Peng. 2021. Learning treatment effects in panels with general intervention patterns. *Advances in Neural Information Processing Systems* **34** 14001–14013.

Farias, Vivek, Andrew Li, Tianyi Peng, Andrew Zheng. 2022. Markovian interference in experiments. *Advances in Neural Information Processing Systems* **35** 535–549.

Farias, Vivek F, Benjamin Van Roy. 2010. Dynamic pricing with a prior on market response. *Operations Research* **58**(1) 16–29.

Geiger, Dan, Thomas Verma, Judea Pearl. 1990. Identifying independence in bayesian networks. *Networks* **20**(5) 507–534.

Gupta, Somit, Ronny Kohavi, Diane Tang, Ya Xu, Reid Andersen, Eytan Bakshy, Niall Cardin, Sumita Chandran, Nanyu Chen, Dominic Coey, et al. 2019. Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explorations Newsletter* **21**(1) 20–35.

Hernán, Miguel A, Tyler J VanderWeele. 2011. Compound treatments and transportability of causal inference. *Epidemiology* **22**(3) 368–377.

Hinnosaar, Marit. 2023. The persistence of healthy behaviors in food purchasing. *Marketing Science* **42**(3) 521–537.

Ho, Teck-Hua, Noah Lim, Sadat Reza, Xiaoyu Xia. 2017. Om forum—causal inference models in operations management. *Manufacturing & Service Operations Management* **19**(4) 509–525.

Hotz, V Joseph, Guido W Imbens, Jacob A Klerman. 2006. Evaluating the differential effects of alternative welfare-to-work training components: A reanalysis of the california gain program. *Journal of Labor Economics* **24**(3).

Hotz, V Joseph, Guido W Imbens, Julie H Mortimer. 2005. Predicting the efficacy of future training programs using past experiences at other locations. *Journal of econometrics* **125**(1-2) 241–270.

Huang, Biwei, Kun Zhang, Mingming Gong, Clark Glymour. 2020. Causal discovery from multiple data sets with non-identical variable sets. *Proceedings of the AAAI conference on artificial intelligence*, vol. 34. 10153–10161.

Huang, Shan, Chen Wang, Yuan Yuan, Jinglong Zhao, Jingjing Zhang. 2023. Estimating effects of long-term treatments. *arXiv preprint arXiv:2308.08152* .

Imbens, Guido, Nathan Kallus, Xiaojie Mao, Yuhao Wang. 2022. Long-term causal inference under persistent confounding via data combination. *arXiv preprint arXiv:2202.07234* .

Imbens, Guido W. 2020. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature* **58**(4) 1129–1179.

Imbens, Guido W, Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jagabathula, Srikanth, Dmitry Mitrofanov, Gustavo Vulcano. 2022. Personalized retail promotions through a directed acyclic graph–based representation of customer preferences. *Operations Research* **70**(2) 641–665.

Johari, Ramesh, Hannah Li, Inessa Liskovich, Gabriel Y Weintraub. 2022. Experimental design in two-sided platforms: An analysis of bias. *Management Science* **68**(10) 7069–7089.

Jung, Yonghan, Iván Díaz, Jin Tian, Elias Bareinboim. 2024. Estimating causal effects identifiable from a combination of observations and experiments. *Advances in Neural Information Processing Systems* **36**.

Kallus, Nathan, Angela Zhou. 2021. Minimax-optimal policy learning under unobserved confounding. *Management Science* **67**(5) 2870–2890.

Lee, Sanghack, Juan D Correa, Elias Bareinboim. 2020. General identifiability with arbitrary surrogate experiments. *Uncertainty in artificial intelligence*. PMLR, 389–398.

Lu, Lauren Xiaoyuan, Jan A Van Mieghem. 2009. Multimarket facility network design with offshoring applications. *Manufacturing & Service Operations Management* **11**(1) 90–108.

Ma, Yu, Kusum L Ailawadi, Dhruv Grewal. 2013. Soda versus cereal and sugar versus fat: drivers of healthful food intake and the impact of diabetes diagnosis. *Journal of Marketing* **77**(3) 101–120.

Maathuis, Marloes H, Diego Colombo. 2015. A generalized back-door criterion. *The Annals of Statistics* 1060–1088.

Markus Kalisch, Martin Mächler, Diego Colombo, Marloes H. Maathuis, Peter Bühlmann. 2012. Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software* **47**(11) 1–26. doi:10.18637/jss.v047.i11.

Meek, Christopher. 1995. Causal inference and causal explanation with background knowledge. *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. 403–410.

Mela, Carl F, Sunil Gupta, Donald R Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing research* **34**(2) 248–261.

Mooij, Joris M, Sara Magliacane, Tom Claassen. 2020. Joint causal inference from multiple contexts. *The Journal of Machine Learning Research* **21**(1) 3919–4026.

Pearl, Judea. 1995. Causal diagrams for empirical research. *Biometrika* 669–688.

Pearl, Judea. 2000. Causality: Models, reasoning and inference. *Cambridge, UK: Cambridge University Press* .

Pearl, Judea, Elias Bareinboim. 2014. External validity: From do-calculus to transportability across populations. *Statistical Science* 579–595.

Perkovi, Emilija, Johannes Textor, Markus Kalisch, Marloes H Maathuis, et al. 2018. Complete graphical characterization and construction of adjustment sets in markov equivalence classes of ancestral graphs. *Journal of Machine Learning Research* **18**(220) 1–62.

Prentice, Ross L. 1989. Surrogate endpoints in clinical trials: definition and operational criteria. *Statistics in medicine* **8**(4) 431–440.

Roberto, Christina A, Ichiro Kawachi. 2015. *Behavioral economics and public health*. Oxford University Press.

Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* **66**(5) 688.

Shannon, Jerry. 2014. Food deserts: Governing obesity in the neoliberal city. *Progress in Human Geography* **38**(2) 248–266.

Shpitser, Ilya, Tyler VanderWeele, James M Robins. 2010. On the validity of covariate adjustment for estimating causal effects. *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*. 527–536.

Singal, Raghav, George Michailidis. 2024. Axiomatic effect propagation in structural causal models. *Journal of Machine Learning Research* **25**(52) 1–71.

Spirtes, Peter, Clark N Glymour, Richard Scheines, David Heckerman, Christopher Meek, Gregory Cooper, Thomas Richardson. 2000. *Causation, prediction, and search*. MIT press.

Spirtes, Peter, Kun Zhang. 2016. Causal discovery and inference: concepts and recent methodological advances. *Applied informatics*, vol. 3. SpringerOpen, 1–28.

Tillman, Robert, David Danks, Clark Glymour. 2008. Integrating locally learned causal structures with overlapping variables. *Advances in Neural Information Processing Systems* **21**.

Tillman, Robert, Peter Spirtes. 2011. Learning equivalence classes of acyclic models with latent and selection variables from multiple datasets with overlapping variables. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. JMLR Workshop and Conference Proceedings, 3–15.

Tong, Simon, Daphne Koller. 2001. Active learning for structure in bayesian networks. *International joint conference on artificial intelligence*, vol. 17. Citeseer, 863–869.

Triantafillou, Sofia, Ioannis Tsamardinos. 2015. Constraint-based causal discovery from multiple interventions over overlapping variable sets. *The Journal of Machine Learning Research* **16**(1) 2147–2205.

Triantafillou, Sofia, Ioannis Tsamardinos, Ioannis Tollis. 2010. Learning causal structure from overlapping variable sets. *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. JMLR Workshop and Conference Proceedings, 860–867.

van der Zander, Benito, Maciej Liśkiewicz, Johannes Textor. 2014. Constructing separators and adjustment sets in ancestral graphs. *Proceedings of the UAI 2014 Conference on Causal Inference: Learning and Prediction-Volume 1274*. 11–24.

Verma, Thomas, Judea Pearl. 1990. Equivalence and synthesis of causal models. *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*. 255–270.

Wang, Guihua, Jun Li, Wallace J Hopp. 2022. An instrumental variable forest approach for detecting heterogeneous treatment effects in observational studies. *Management Science* **68**(5) 3399–3418.

Xiong, Ruoxuan, Susan Athey, Mohsen Bayati, Guido Imbens. 2024. Optimal experimental design for staggered rollouts. *Management Science* **70**(8) 5317–5336.

Yang, Jeremy, Dean Eckles, Paramveer Dhillon, Sinan Aral. 2023. Targeting for long-term outcomes. *Management Science* .

Ye, Zikun, Zhiqi Zhang, Dennis Zhang, Heng Zhang, Renyu Philip Zhang. 2023. Deep-learning-based causal inference for large-scale combinatorial experiments: Theory and empirical evidence. *Available at SSRN 4375327* .

Zhang, Kun, Biwei Huang, Jiji Zhang, Clark Glymour, Bernhard Schölkopf. 2017. Causal discovery from nonstationary/heterogeneous data: Skeleton estimation and orientation determination. *IJCAI: Proceedings of the Conference*, vol. 2017. NIH Public Access, 1347.

Zhou, Zhengyuan, Susan Athey, Stefan Wager. 2023. Offline multi-action policy learning: Generalization and optimization. *Operations Research* **71**(1) 148–183.

Appendix

A. Meek Rules Illustration

The Meek rules provide a systematic method for orienting undirected edges within a partially directed acyclic graph to achieve a fully directed acyclic graph without introducing new unshielded collider or cycles (Meek 1995). Figure 9 demonstrates these rules visually, showcasing how they apply in different configurations of a graph.

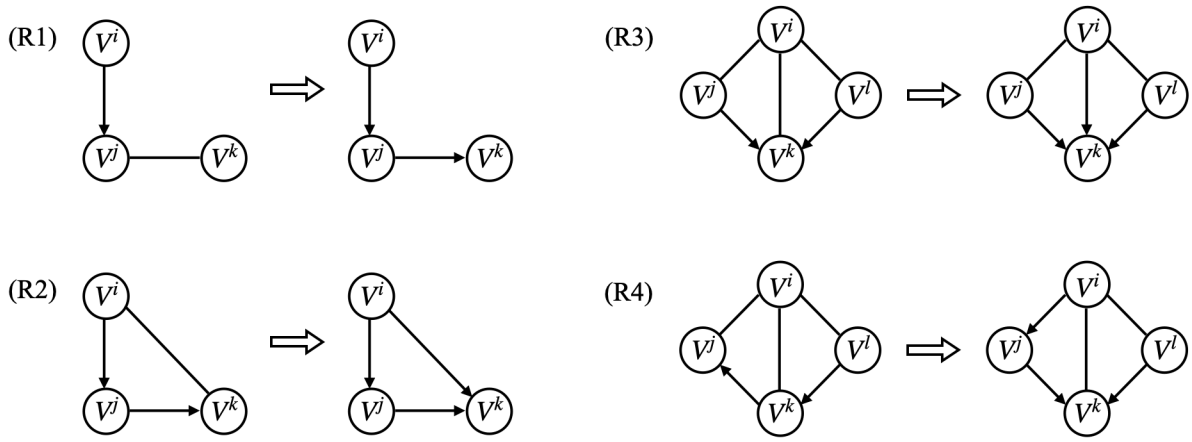

Figure 9 **Meek rules illustrated.**

B. Rules of Do-Calculus

Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, and $\mathbf{W}$ be arbitrary disjoint sets of nodes in a causal DAG $\mathcal{G}$ and let P be the probability distribution induced by graph $\mathcal{G}$. According to Theorem 3.4.1 in Pearl (2000), the following rules apply to any disjoint subsets of variables $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, and $\mathbf{W}$:

1. **Rule 1 (Insertion/deletion of observations):**

$$P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \text{ if } (\mathbf{Y} \perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}}}}.$$

2. **Rule 2 (Action/observation exchange):**

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}) \text{ if } (\mathbf{Y} \perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}}.$$

3. **Rule 3 (Insertion/deletion of actions):**

$$P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \text{ if } (\mathbf{Y} \perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{\mathcal{G}_{\overline{\mathbf{X},\mathbf{Z}(\mathbf{W})}}}.$$

where $\mathcal{G}_{\overline{\mathbf{X}}}$ denotes the graph obtained by deleting all incoming arrows to nodes in $\mathbf{X}$ from $\mathcal{G}$, $\mathcal{G}_{\underline{\mathbf{X}}}$ denotes the graph obtained by deleting all outgoing arrows from nodes in $\mathbf{X}$ from $\mathcal{G}$, and $\mathbf{Z}(\mathbf{W})$ is the set of $\mathbf{Z}$-nodes that are not ancestors of any $\mathbf{W}$-nodes in $\mathcal{G}_{\overline{\mathbf{X}}}$.

C. Proofs

LEMMA 1. *In a directed acyclic graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, if V_1 is adjacent to V_2, and V_2 is adjacent to V_3, and V_1 is not adjacent to V_3, then the edges are oriented as $V_1 \rightarrow V_2 \leftarrow V_3$ if and only for every subset $\mathbf{C}$ of V, V_1 is d-connected to V_3 given $\{V_2\} \cup C \setminus \{V_1, V_3\}$.*

Proof of Lemma 1.The proof follows from Lemma 5.1.2 used in the proof of Theorem 5.1 in Spirtes et al. (2000).

LEMMA 2. *In a directed acyclic graph $\mathcal{G} = (\mathbf{V}, \mathbf{E})$, if V_1 is adjacent to V_2, and V_2 is adjacent to V_3, and V_1 is not adjacent to V_3, then either V_2 is in every set of variables that d-separates V_1 and V_3, or it is in no set of variables that d-separates V_1 and V_3.*

Proof of Lemma 2.The proof follows from Lemma 5.1.3 used in the proof of Theorem 5.1 in Spirtes et al. (2000).

LEMMA 3. *Two DAGs are equivalent if and only if they share the same skeleton and same v-structures.*

Proof of Lemma 3. The proof of this lemma can be found in Verma and Pearl (1990).

Proof of Theorem 2. We prove this in two steps. First, we show that a graph $\mathcal{G} \in \mathbf{G}$ returned by the COMB-PC algorithm shares the same skeleton with the true underlying DAG $\mathcal{G}^*$. Second, we show that the v-structures on any $\mathcal{G} \in \mathbf{G}$ and $\mathcal{G}^*$ are identical. This approach is taken in light of Lemma 3, which states that two DAGs are equivalent if and only if they share the same skeleton and the same unshielded colliders, thus guiding our proof structure towards establishing Markov equivalence between $\mathcal{G} \in \mathbf{G}$ and $\mathcal{G}^*$.

Part 1. We start by noting that all the graphs in $\mathbf{G}$ returned by phase 2 of COMB-PC algorithm have the same skeleton as the undirected graph $\mathcal{G}_1$ returned in phase 1. Therefore, it would be sufficient to show $\mathcal{G}_1$ is identical to the skeleton of the true underlying DAG $\mathcal{G}^*$. Let's select $V_i, V_j \in \mathbf{V}^E$ such that V_i and V_j are *not neighbors* in $\mathcal{G}_1$. Since V_i and V_j are *not neighbors* in $\mathcal{G}_1$, then there must exist a conditioning set $\mathbf{C}^* \in \mathbf{C}^E_{V_i,V_j}$ such that $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}^*$ due to the construction of Algorithm 1. Note that Assumptions 1, 2, and 7 ensures that we have $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}^*$ reflects the d-separation relations in the true graph. Hence V_1 and V_2 must be d-separated given $\mathbf{C}^*$ in $\mathcal{G}^*$. Hence V_i and V_j cannot be neighbors in $\mathcal{G}^*$ by Definitions 1 and 2. Now let's select $V_i, V_j \in \mathbf{V}^E$ such that V_i and V_j are *not neighbors* in $\mathcal{G}^*$. Then there must exists a set $\mathbf{C}^{**} \in \mathbf{C}^E_{V_i,V_j}$ that d-separates V_i and V_j in $\mathcal{G}^*$. By definition, $|\mathbf{C}^{**}| \leq |\mathbf{V}| - 2$ where $|\cdot|$ gives the size of a set. Since Y can only be a collider in $\mathcal{G}^*$ by Assumption 5(*ii*), we further have $|\mathbf{C}^{**}| \leq |\mathbf{V} \setminus \{Y\}| - 2 = |\mathbf{V}| - 3$. Note that we assumed we have the perfect conditional independence information, we must have $V_i \perp\!\!\!\perp_E V_j \mid \mathbf{C}^{**}$. The first phase of COMB-PC algorithm iterates until it finds a conditioning set that makes V_i and

V_j independent with size less than or equal to $|V|-3$. Hence it eventually reaches the set $\mathbf{C}^{**}$ and removes the edge V_i—V_j from U (if the edge V_i—V_j is not removed earlier). With this, we establish V_i and V_j are not neighbors in $\mathcal{G}_1$, which implies V_i and V_j are not neighbors in any of the graphs in $\mathbf{G}$ as well. Similarly, we can show Y and $V_i \in \mathbf{V} \setminus X$ are not neighbors in $\mathcal{G}_1$ if and only if they are not neighbors in $\mathcal{G}^*$. Finally, let's consider the pair W and Y. Note that W and Y are not adjacent in $\mathcal{G}^*$ by Assumption 5(i) and they are not adjacent in any of the graphs in $\mathbf{G}$ as an edge between W and Y is never included in $\mathbf{U}$. Note that both the experimental and observational samples have the same distribution by Assumption 6. That's why the adjacency relations found using the observational data for Y will be valid in the experimental data as well. With this, we prove that the skeleton of $\mathcal{G}_1$ returned by the first phase of the COMB-PC algorithm is identical to the skeleton of the true underlying DAG $\mathcal{G}^*$, which implies $\mathcal{G} \in \mathbf{G}$ and $\mathcal{G}^*$ shares the same skeleton.

Part 2. We next show for all $\mathcal{G} \in \mathbf{G}$, $\mathcal{G}$ has the same unshielded colliders with $\mathcal{G}^*$. Since $\mathbf{G}$ stores all DAGs within the Markov equivalence class characterized by $\mathcal{G}_2 = (\mathbf{V}, \mathbf{M})$, any graph $\mathcal{G} \in \mathbf{G}$ must share the same unshielded colliders as $\mathcal{G}_2$ by Lemma 3. Step 1. Here we show every unshielded collider in $\mathcal{G}_2$ is also in $\mathcal{G}^*$. Without loss of generality, let's select any $V_i, V_j, V_k \in \mathbf{V}$ where V_i, V_j, V_k form an unshielded collider on graph $\mathcal{G}_2$. Note for any $V_i, V_j, V_k \in \mathbf{V}$ where V_i, V_j, V_k form an unshielded collider on graph $\mathcal{G}_2$, we have $V_j \neq W$ as we make sure there are no incoming edges to W in $\mathcal{G}_2$ in the treatment randomization step in phase 2. By definition of an unshielded collider, we have V_i and V_k are not neighbors in $\mathcal{G}_2$ and $(V_i \to V_j)$ and $(V_j \leftarrow V_k)$ are in $\mathcal{G}_2$. Step 1A. Let's first focus on the case where $Y \notin \{V_i, V_j, V_k\}$, i.e., $\{V_i, V_j, V_k\} \subseteq \mathbf{V}^E$. Note that all unshielded colliders over variables $\mathbf{V}^E$ will be oriented through the dependence relations only as Meek's rule are guaranteed to not to create further unshielded colliders (Meek 1995). Since we have $(V_i \to V_j)$ and $(V_j \leftarrow V_k)$ are in $\mathcal{G}_2$, we must have $V_j \notin SepSet_{V_i V_k}$ by construction of the COMB-PC algorithm, which implies $V_i \not\perp\!\!\!\perp_E V_k \mid V_j$. Since we assumed perfect conditional independence information by Assumption 7, we must have V_i and V_k d-connected with respect to V_j in $\mathcal{G}^*$. Note that we already established that $\mathcal{G}_2$ and $\mathcal{G}^*$ share the same skeleton in the previous part of the proof. This implies V_i and V_j and V_j and V_k are *adjacent* and V_i and V_k are *not adjacent* in $\mathcal{G}^*$. Then by Lemmas 1 and 2, $V_i \to V_j \leftarrow V_k$ must be present in the true underlying DAG $\mathcal{G}^*$ forming an unshielded collider. Step 1B. Now we focus on the case where $Y \in \{V_i, V_j, V_k\}$. Note that we make sure there are no outgoing edges from Y in $\mathcal{G}_2$ in the long-term outcome integration step of phase 2 of the COMB-PC Algorithm. Hence for any V_i, V_j, V_k that form an unshielded collider on graph $\mathcal{G}_2$ such that $Y \in \{V_i, V_j, V_k\}$, we must have $V_j = Y$, i.e., $(V_i \to Y)$ and $(Y \leftarrow V_k)$. Since we already established $\mathcal{G}_2$ and $\mathcal{G}^*$ share the same skeleton, the skeleton of $\mathcal{G}^*$ must include V_i—Y and V_k—Y. Then by Assumption 5(ii), we must have $(V_i \to Y)$ and $(Y \leftarrow V_k)$ in $\mathcal{G}^*$. With this, we show V_i, V_j, V_k form an unshielded collider on the true underlying DAG $\mathcal{G}^*$. Note that Meek rules make sure no new unshielded colliders are introduced (Meek 1995). With this, we

prove every unshielded collider in $\mathcal{G}_2$ is also in $\mathcal{G}^*$. Step 2. Here we show every unshielded collider in $\mathcal{G}^*$ is also in $\mathcal{G}_2$. Without loss of generality, let's now select $V_{i'}, V_{j'}, V_{k'} \in \mathbf{V}$ where $V_{i'}, V_{j'}, V_{k'}$ form an unshielded collider on graph $\mathcal{G}^*$. This means $V_{i'}$ and $V_{k'}$ are not neighbors in $\mathcal{G}^*$ and we have $(V_{i'} \to V_{j'})$ and $(V_{j'} \leftarrow V_{k'})$ in $\mathcal{G}^*$. Step 2A. Let's first focus on the case where $Y \notin \{V_i, V_j, V_k\}$, i.e., $\{V_i, V_j, V_k\} \subseteq \mathbf{V}^E$. Since we showed $\mathcal{G}^*$ and $\mathcal{G}_2$ share the same skeleton, we must have $V_{i'}$ and $V_{j'}$ and $V_{j'}$ and $V_{k'}$ neighbors in graph $\mathcal{G}_2$, and $V_{i'}$ and $V_{k'}$ are not neighbors in $\mathcal{G}_2$. Note that we must have $V_{j'} \neq W$ by Assumption 4. Note that $V_{i'}$ and $V_{k'}$ are d-connected with respect to $V_{j'}$ in graph $\mathcal{G}^*$ as $V_{i'} \to V_{j'}$ and $V_{j'} \leftarrow V_{k'}$. Since we assumed we have access to perfect conditional independence information, we must have $V_{j'} \notin SepSet_{V_{i'}V_{k'}}$. Then phase 2 of the COMB-PC algorithm ensures that we have $(V_{i'} \to V_{j'})$ and $(V_{j'} \leftarrow V_{k'})$ in $\mathcal{G}_2$. With this we show $V_{i'}, V_{j'}, V_{k'}$ forms an unshielded collider on $\mathcal{G}_2$. Step 2B. Now we focus on the case where $Y \in \{V_i, V_j, V_k\}$. By Assumption 5(ii), we must have $V_{j'} = Y$ in $\mathcal{G}^*$. Note that since we already established $\mathcal{G}^*$ and $\mathcal{G}_2$ share the same skeleton, $V_{i'}$ and $V_{k'}$ are also neighbors with Y in $\mathcal{G}_2$. Then by the long-term outcome integration step of phase 2 of the COMB-PC algorithm, we must have $V_{i'} \to Y$ and $V_{k'} \to Y$ in $\mathcal{G}_2$. Hence V_i, V_j, V_k form an unshielded collider on graph $\mathcal{G}_2$. With this, we prove every unshielded collider in $\mathcal{G}^*$ is also in $\mathcal{G}_2$. □

Proof of Proposition 1. Under Assumption 4, there are no incoming edges to treatment W in graph $\mathcal{G}$ (Pearl 2000). Hence there are no backdoor paths between W and $\mathbf{S}_\mathcal{G}$ in $\mathcal{G}$ and we have $P(\mathbf{S}_\mathcal{G} = \mathbf{s}_\mathcal{G} | D = E, do(W = w)) = P(\mathbf{S}_\mathcal{G} = \mathbf{s}_\mathcal{G} | D = E, W = w)$. □

Proof of Proposition 2. We prove Proposition 2 by showing $\mathbf{Z}_\mathcal{G}$ satisfies the backdoor criterion relative to every variable in $\mathbf{S}_\mathcal{G}$ and the outcome Y in the $\mathcal{G}$ using Definition 3. We will consider conditions (i) and (ii) of Definition 3 separately. Condition (i). We prove this by contradiction. Suppose there exists a $Z \in \mathbf{Z}_\mathcal{G}$ where Z is a descendant of some $S \in \mathbf{S}_\mathcal{G}$. By construction of Algorithm 3, all $S \in \mathbf{S}$ lies on a directed path from W to Y. This implies there exists a directed path from W to S in the graph $\mathcal{G}$. Since we assumed Z is a descendant of S, there must exist a directed path from W to Z in $\mathcal{G}$. Note again by construction of Algorithm 3, there must exist a backdoor path p' between some $S' \in \mathbf{S}_\mathcal{G}$ and Y where $\psi(p') = Z$. Hence there must exist an edge between Z and Y in $\mathcal{G}$, i.e., $Z \in \mathcal{N}(Y)$. By Step 3 of Algorithm 2 and Assumption 5, we must have the edge $Z \to Y$ in $\mathcal{G}$. So far we established there must exist a directed path from W to Z and we have the edge $Z \to Y$ in $\mathcal{G}$. Together they imply there must exist a directed path p from W to Y where Z is the second-to-last element, i.e., $\psi(p) = Z$. Then by Step 1 of Algorithm 3, Z must be in $\mathbf{S}_\mathcal{G}$. However for a backdoor path p' between any $S' \in \mathbf{S}_\mathcal{G}$ where $\psi(p') = Z$, we have $Z \in noncolliders(p')$. Therefore, it follows that for every backdoor path p', the set intersection $noncolliders(p') \cap (\mathbf{S}_\mathcal{G} \setminus \{S\})$ is non-empty because Z is contained within $\mathbf{S}_\mathcal{G} \setminus \{S\}$. It is therefore impossible for Algorithm 3 to include Z in $\mathbf{Z}_\mathcal{G}$. This result contradicts the assumption that Z is in $\mathbf{Z}_\mathcal{G}$. Condition (ii). Here we show every backdoor path between any $S \in \mathbf{S}_\mathcal{G}$ and Y is blocked with respect to the set $\mathbf{Z}_\mathcal{G} \cup \mathbf{S}_\mathcal{G} \setminus S$. Note that

the definition of blocked paths is given in Definition 1. Algorithm 4 iterates over all $S \in \mathbf{S}_{\mathcal{G}}$ and over all backdoor paths between S and Y. Without loss of generality, let's fix $S^* \in \mathbf{S}_{\mathcal{G}}$ and $p'' \in B_{S^*Y}$. Next,we will show p'' is blocked with respect to the set $\mathbf{Z}_{\mathcal{G}} \cup \mathbf{S}_{\mathcal{G}} \setminus S^*$. Suppose the if condition in Step 1 of Algorithm 4 doesn't hold, i.e., $noncolliders(p'') \cap \mathbf{S}_{\mathcal{G}} \setminus S^* \neq \emptyset$. Then, p'' is blocked with respect to $\mathbf{Z}_{\mathcal{G}} \cup \mathbf{S}_{\mathcal{G}} \setminus S^*$ by Definition 1. Now suppose the if condition in Step 1 of Algorithm 4 holds, i.e., $noncolliders(p'') \cap \mathbf{S}_{\mathcal{G}} \setminus S^* = \emptyset$. Then $\mathbf{Z}_{\mathcal{G}}$ includes the second-to-last node on path p'' by Algorithm 3. Let Z^* be the second-to-last node on path p'', i.e., $\psi(p'') = Z^*$. This implies that $Z^* \in \mathcal{N}(Y)$ and by Algorithm 2 and Assumption 5, we have $Z^* \to Y$ in $\mathcal{G}$ and do not have $Z^* \leftarrow Y$ in $\mathcal{G}$. Hence Z^* is a non-collider on path p''. Then, p'' is blocked with respect to $\mathbf{Z}_{\mathcal{G}} \cup \mathbf{S}_{\mathcal{G}} \setminus S^*$ by Definition 1. Hence we showed $\mathbf{Z}_{\mathcal{G}}$ satisfies the back-door condition relative to $(\mathbf{S}_{\mathcal{G}}, \mathbf{Y})$ in $\mathcal{G}$.

LEMMA 4. *Consider a graph $\mathcal{G}$ that satisfies Assumptions 4 and 5. Let $\mathbf{S}_{\mathcal{G}}$ and $\mathbf{Z}_{\mathcal{G}}$ denote the sets obtained by applying Algorithms 3 and 4, respectively, on graph $\mathcal{G}$. Then we have $(\mathbf{Y} \perp \mathbf{W} | \mathbf{S}_{\mathcal{G}}, \mathbf{Z}_{\mathcal{G}})_{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$ where $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ is the graph obtained after removing all the incoming edges to $\mathbf{S}_{\mathcal{G}}$ from $\mathcal{G}$.*

Proof of Lemma 4. Let $P_{WY}^{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$ store all the paths from W to Y and $D_{WY}^{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$ store only the *directed* paths from W to Y in graph $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$. Let $nodes(p)$ store the nodes on a path p. Notice that by construction of Algorithm 3, $D_{WY}^{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}} = \emptyset$. Without loss of generality, let's select a path $p' \in P_{WY}^{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$. We will show that path p' is blocked with respect to the sets $\mathbf{S}_{\mathcal{G}}, \mathbf{Z}_{\mathcal{G}}$. Since $\mathcal{G}$ satisfies Assumption 4, i.e. there are no incoming edges to W, and since $D_{WY}^{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}} = \emptyset$, there must exist a collider on path p'. Let $col^1(p')$ be the first collider on path p'. Part 1. We first show $col^1(p') \notin \mathbf{S}_{\mathcal{G}} \cup \mathbf{Z}_{\mathcal{G}}$. Since there are no incoming edges to $\mathbf{S}_{\mathcal{G}}$ in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ by definition, we have $col^1(p') \notin \mathbf{S}_{\mathcal{G}}$. Now suppose $col^1(p') \in \mathbf{Z}_{\mathcal{G}}$. Let $\overline{p'}$ represent the sub-path of p' from W to $col^1(p')$. Since $col^1(p')$ is the first collider on path p' and since there are no incoming edges to W in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ as a result of Assumption 4, then $\overline{p'}$ is a directed path from W to $col^1(p')$. Since we assumed $col^1(p') \in \mathbf{Z}_{\mathcal{G}}$ and since the edge $Z \to Y$ is included in the graph $\mathcal{G}$ for all $Z \in \mathbf{Z}_{\mathcal{G}}$ by construction of Algorithm 4, we must have $col^1(p') \to Y$ included in graph $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$. However, this implies that there must exist a directed path from W to Y where $col^1(p')$ is the second-to-last element. Consequently, it must be the case that $col^1(p') \in \mathbf{S}_{\mathcal{G}}$ as a result of Algorithm 3. Therefore p' cannot exist in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$. This leads to a contradiction. Hence we must have $col^1(p') \notin \mathbf{Z}_{\mathcal{G}}$. Part 2. Next, we demonstrate that the set $\mathbf{S}_{\mathcal{G}} \cup \mathbf{Z}_{\mathcal{G}}$ does not contain any descendants of $col^1(p')$ within the graph $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$. Note that none of the nodes in $\mathbf{S}_{\mathcal{G}}$ can be a descendant of $col^1(p')$ in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ as $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ doesn't include any incoming edges to $\mathbf{S}_{\mathcal{G}}$. Next we show $\mathbf{Z}_{\mathcal{G}}$ doesn't include a descendant of $col^1(p')$. We prove this by contradiction. Suppose there exists $Z' \in \mathbf{Z}_{\mathcal{G}}$ where Z' is a descendant of $col^1(p')$. Note that $\mathcal{G}$ must include the edge $Z' \to Y$ by construction of Algorithm 4, which implies $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$ includes the edge $Z' \to Y$. Since we assumed Z' is a descendant of $col^1(p')$ and since $\overline{p'}$ is a directed path from W to $col^1(p')$ in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$, there exists a directed path from W to Y where Z' is the second-to-last element.

Hence we must have $Z' \in \mathbf{S}_{\mathcal{G}}$, however, this conflicts with the steps of Algorithm 4. Therefore Z' cannot be a descendant of $col^1(p')$.

We thus far demonstrated that there exists a collider $col^1(p')$ on path p' where $col^1(p') \notin \mathbf{S}_{\mathcal{G}} \cup \mathbf{Z}_{\mathcal{G}}$ and $\mathbf{S}_{\mathcal{G}} \cup \mathbf{Z}_{\mathcal{G}}$ doesn't include any descendants of $col^1(p')$. Then by Definition 1, path p' between W and Y is blocked. This will hold for all paths between W and Y in $\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}$. Hence by Definition 2, we have $(\mathbf{Y} \perp \mathbf{W} | \mathbf{S}_{\mathcal{G}}, \mathbf{Z}_{\mathcal{G}})_{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$. □

Proof of Proposition 3. Note that this proof utilizes the rules of do-calculus as outlined by Pearl (2000). For detailed information on these rules, please refer to Appendix B. We will use the following notation in this proof: $\mathcal{G}_{\overline{X}}$ denotes the graph obtained by deleting all incoming arrows to nodes in $\mathbf{X}$ from $\mathcal{G}$. Similarly, $\mathcal{G}_{\underline{\mathbf{X}}}$ denotes the graph obtained by deleting all outgoing arrows from nodes in X from $\mathcal{G}$.

The proof proceeds as follows:

$$P(Y = y \mid D = E, \text{do}(W = w)) = \sum_{\mathbf{z}_{\mathcal{G}} \in \mathcal{Z}_{\mathcal{G}}} P(Y = y \mid D = E, \text{do}(W = w), \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}) P(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}} \mid D = E, \text{do}(W = w)). \tag{8a}$$

Equation (8a) holds by the law of total probability. Again by the law of total probability we have

$$\begin{aligned} P(Y = y \mid D = E, \text{do}(W = w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} & P(Y = y \mid D = E, \text{do}(W = w), \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}, \mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}}) \\ & \times P(\mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}} \mid D = E, \text{do}(W = w), \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}) \\ & \times P(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}} \mid D = E, \text{do}(W = w)). \end{aligned} \tag{8b}$$

Since $\mathbf{Z}_{\mathcal{G}}$ blocks all the backdoor paths from $\mathbf{S}_{\mathcal{G}}$ to Y by Proposition 2, we have $(\mathbf{Y} \perp \mathbf{S}_{\mathcal{G}} | W, \mathbf{Z}_{\mathcal{G}})_{\mathcal{G}_{\overline{W}\underline{\mathbf{S}}_{\mathcal{G}}}}$ as . Then by rule 2 of do-calculus, we have:

$$\begin{aligned} P(Y = y \mid D = E, \text{do}(W = w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} & P(Y = y \mid D = E, \text{do}(W = w), \text{do}(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}), \mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}}) \\ & \times P(\mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}} \mid D = E, \text{do}(W = w), \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}) \\ & \times P(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}} \mid D = E, \text{do}(W = w)). \end{aligned} \tag{8c}$$

Next we can replace $\text{do}(W = w)$ with $W = w$ by Assumption 4.

$$\begin{aligned} P(Y = y \mid D = E, \text{do}(W = w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} & P(Y = y \mid D = E, W = w, \text{do}(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}), \mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}}) \\ & \times P(\mathbf{S}_{\mathcal{G}} = \mathbf{s}_{\mathcal{G}} \mid D = E, W = w, \mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}}) \\ & \times P(\mathbf{Z}_{\mathcal{G}} = \mathbf{z}_{\mathcal{G}} \mid D = E, W = w). \end{aligned} \tag{8d}$$

Note that we have $(Y \perp \mathbf{W} | \mathbf{S}_{\mathcal{G}}, \mathbf{Z}_{\mathcal{G}})_{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$ by Lemma 4. Then by rule 1 of do-calculus, we have:

$$P(Y=y \mid D=E, \mathrm{do}(W=w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} P(Y=y \mid D=E, \mathrm{do}(\mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}), \mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}}) \times P(\mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}} \mid D=E, W=w, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}) \times P(\mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}} \mid D=E, W=w). \tag{8e}$$

Note that we have $(Y \perp \mathbf{S}_{\mathcal{G}} | \mathbf{Z}_{\mathcal{G}})_{\mathcal{G}_{\overline{\mathbf{S}}_{\mathcal{G}}}}$ as Proposition 2 shows that $\mathbf{Z}_{\mathcal{G}}$ satisfies the backdoor criterion given in Definition 3. Let's introduce a variable $\mathbf{X}$ and set $\mathbf{X} = \emptyset$ to invoke rule 2 of do-calculus. Then, we have :

$$P(Y=y \mid D=E, \mathrm{do}(W=w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} P(Y=y \mid D=E, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}, \mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}}) \times P(\mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}} \mid D=E, W=w, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}) \times P(\mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}} \mid D=E, W=w). \tag{8f}$$

Furthermore, the proof of Lemma 4 shows that all paths between $\mathbf{Z}_{\mathcal{G}}$ and W in $\mathcal{G}$ must include a collider. Hence we have $(\mathbf{Z}_{\mathcal{G}} \perp W)_{\mathcal{G}_{\overline{W}}}$. Then by the rule 1 of do-calculus, we have:

$$P(Y=y \mid D=E, \mathrm{do}(W=w)) = \sum_{z \in \mathcal{Z}} \sum_{s \in \mathbf{S}} P(Y=y \mid D=E, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}, \mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}}) \times P(\mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}} \mid D=E, W=w, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}) \times P(\mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}} \mid D=E). \tag{8g}$$

Finally, by Assumption 6 we have:

$$P(Y=y \mid D=E, \mathrm{do}(W=w)) = \sum_{\mathbf{z} \in \mathcal{Z}_{\mathcal{G}}} \sum_{\mathbf{s}_{\mathcal{G}} \in \mathbf{S}_{\mathcal{G}}} P(Y=y \mid D=O, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}, \mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}}) \times P(\mathbf{S}_{\mathcal{G}}=\mathbf{s}_{\mathcal{G}} \mid D=E, W=w, \mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}}) \times P(\mathbf{Z}_{\mathcal{G}}=\mathbf{z}_{\mathcal{G}} \mid D=E). \square \tag{8h}$$

Proof of Theorem 3. Under Assumptions 1–7, Theorem 2 proves that the true graph $\mathcal{G}^*$ is included in $\mathbf{G}$ returned by the COMB-PC algorithm. Then, by Proposition 3, the true average treatment effect $\tau_{\mathcal{G}^*}$ must be included in $\tau_{\mathcal{G}^*} \in \{\tau_{\mathcal{G}} \mid \mathcal{G} \in \mathbf{G}\}$ as $\mathcal{G}^* \in \mathcal{G}$. $\square$