

Jeziorski, Przemyslaw; Leng, Dingzhe; Seiler, Stephan

Working Paper

Causal Inference with Endogenous Price Response

CESifo Working Paper, No. 11898

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Jeziorski, Przemyslaw; Leng, Dingzhe; Seiler, Stephan (2025) : Causal Inference with Endogenous Price Response, CESifo Working Paper, No. 11898, CESifo GmbH, Munich

This Version is available at:

<https://hdl.handle.net/10419/320119>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Causal Inference with Endogenous Price Response

Przemyslaw Jeziorski, Dingzhe Leng, Stephan Seiler

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.ifo.de/en/cesifo/publications/cesifo-working-papers>

An electronic version of the paper may be downloaded

· from the SSRN website: www.SSRN.com

· from the RePEc website: www.RePEc.org

· from the CESifo website: <https://www.ifo.de/en/cesifo/publications/cesifo-working-papers>

Causal Inference With Endogenous Price Response

Przemyslaw Jeziorski

UC Berkeley

Dingzhe Leng

UC Berkeley

Stephan Seiler

Imperial College London
& CEPR

This draft: May 10, 2025

We study the estimation of causal treatment effects on demand when treatment is randomly assigned but prices adjust in response to treatment. We show that regressions of demand on treatment or on treatment and price lead to biased estimates of the direct treatment effect. The bias in both cases depends on the correlation of price with treatment and points in the same direction. In most cases including an endogenous price control reduces bias but does not remove it. We show how to test whether bias from an endogenous price response arises and how to recover an unbiased treatment effect (holding price constant) using a price instrument. We apply our approach to the estimation of the impact of feature advertising across several product categories using supermarket scanner data and show that the bias when not instrumenting for price can be substantial.

Keywords: Causal Inference, Endogeneity, Endogenous Controls, Instrumental Variables

1 Introduction

A common research question in economics and marketing is how a specific treatment, such as advertising, customer word-of-mouth, or a nutritional label, affects demand. Although not always explicitly stated, the goal is typically to identify the impact of the treatment on demand *holding price constant*; that is, to estimate the shift in the demand curve caused by the intervention. The estimation of this direct treatment effect is often complicated by the fact that a shift in demand caused by the treatment may lead to price adjustment among treated units. For example, an online marketplace might run an A/B test that randomly improves the product detail pages for certain third party sellers. Such an enhancement increases demand for treated products, which in most cases will prompt the treated sellers to raise prices. As a consequence, the measured change in demand is smaller than the underlying shift in the demand curve because higher prices offset the direct treatment effect.¹ To account for such endogenous price adjustments, many studies attempt to isolate the direct treatment effect by using model specifications that control for price. However, as we demonstrate, using price as a control variable results in a biased estimate of the direct treatment effects, even if treatment is exogenous. This is because price is typically correlated both with the treatment and the error term due to the post-treatment endogenous price adjustment. In this paper we analyze the bias originating from endogenous price adjustments, provide testable conditions for the existence of such bias, and show how to obtain unbiased estimates of the direct causal effect of treatment.

Our analysis proceeds as follows. In Section 2, we assume random treatment assignment, a linear demand function, and a reduced-form pricing equation to derive expressions for the bias in the treatment effect estimate in two scenarios: when price is omitted from the regression, and when price is included as an endogenous control variable. We derive the formula for the direction and relative magnitude of the bias in both situations. We also show that the bias in the treatment coefficient, when using an endogenous price control, can be expressed as a function of the bias in the price coefficient. We then demonstrate how to obtain an unbiased estimate using price instruments and show that no correlation between treatment and price is a sufficient condition for the absence of bias. Importantly, this condition is testable. In Section 3, we illustrate these theoretical findings by conducting a series of numerical simulations. We start with linear demand and supply equations, and then expand our analysis to non-linear settings common in structural demand and supply frameworks. Finally, in Section 4, we use supermarket scanner data across multiple product categories to estimate the impact of feature advertising on demand, controlling for prices. We show that failing to use a valid instrument when controlling for prices can lead to substantial bias, overestimating the effect of feature ads by a factor of 3 to 4 in some categories.

Our research contributes to a large literature in marketing and economics that seeks to estimate

¹In some settings, the equilibrium effect on demand may be of interest. We take as given that the research objective is to estimate the *direct* causal effect when holding price constant. As we show below, many papers explicitly control for price, indicating that their aim is to isolate the direct impact of a given treatment rather than capturing equilibrium market response.

the causal impact of firm or policy interventions on demand. Many studies include price as an endogenous control variable. A non-exhaustive list of papers using such a specification includes the sales impact of feature advertising (Seiler and Yao, 2017), Super Bowl advertising (Hartmann and Klapper, 2018), TV advertising for consumer-packages goods (Shapiro et al., 2021), the effect of income and wealth on private label demand (Dubé et al., 2018), the effect of product labels such as nutritional labels (Barahona et al., 2023),² and “Made in USA” labels (Kong and Rao, 2021).³ Price endogeneity is generally considered a major concern in studies focusing on estimating price elasticities, such as demand estimation for the purpose of informing firms’ optimal pricing strategies or merger evaluations. However, when price is merely used as a control, these concerns are often perceived as less important. For example Kong and Rao (2021) state “The endogeneity of price [...] does not affect our main coefficient of interest [...] because the exogeneity of treatment still holds conditional on price ...”. As our analysis shows, this claim is only correct under specific conditions. In most cases a biased price coefficients leads to a biased treatment coefficient, even when treatment is randomly assigned. Few papers implement the test we propose for endogenous price responses, i.e., evaluate whether treatment and price are correlated, possibly conditional on other control variables. Among the papers mentioned above, Shapiro et al. (2021) implement such a test and find that price and advertising are uncorrelated and hence an endogenous price response does not pose a concern.⁴

Our setting fits within a broader causal inference literature that analyzes regressions with endogenous control variables. Rosenbaum (1984) shows that controlling for an endogenous regressor, which he refers to as a “post-treatment variable”, can yield a biased estimate of the treatment effect even when treatment is randomized. Similarly, Angrist and Pischke (2009) define “bad controls” as variables that are “themselves outcome variables” – in our case, prices. Perils of controlling for post-treatment variables has been discussed across various fields, such as, medicine (Streiner, 2016), political science (Montgomery et al., 2018), sociology (Cinelli et al., 2024), and organizational research (Mändli and Rönkkö, 2025). Most of these papers highlight biases when including endogenous controls and recommend not including them and thus estimating only the joint effect. Our paper goes beyond such an analysis by deriving explicit expressions for the bias arising from endogenous controls, showing how to test for endogenous adjustments, and proposing a solution that allows the researcher to estimate direct and indirect effects. We believe such an analysis is important for settings like the one we analyze where isolating only the direct effect of treatment (the impact on demand holding price constant) is often of interest to the researcher.

Our paper is also closely related to the mediation analysis literature, where the focus is explicitly on distinguishing direct and indirect effect. Researchers frequently conduct “path analysis” by

²In their reduced-form analysis, Barahona et al. (2023) regress demand on label dummies and price without instrumenting the latter. In their structural model they do use price instruments.

³The broader practice of using endogenous post-treatment variables as controls is also prevalent in other fields. Acharya et al. (2016) survey the empirical political science literature and find that two-thirds of papers use such controls.

⁴Dubé et al., 2018 regress price on macroeconomic indicators and find that some of them have a significant effect on prices.

controlling for intermediate variables along the causal pathway. They subsequently check if the original effect disappears to establish whether a direct causal relationship remains (Hayes, 2017). Several papers across different disciplines highlight that mediation without random variation in the mediator leads to biased estimates (e.g. Acharya et al. (2016), Judd and Kenny (1981), Petersen et al. (2006)). Several solutions have been proposed. Acharya et al. (2016) propose a sequential estimator that is valid based on two separate conditional independence assumption with regards to the treatment and the intermediate variable. Imai et al. (2011) similarly propose a sequential ignorability assumption which assumes that the mediator (price in our setting) is not affected by any variable that also affects the outcome (except for the treatment itself) conditional on a set of pre-treatment control variables. This assumption is unlikely to be fulfilled in our setting because unobserved factors that impact demand will also likely impact prices. Bullock and Green (2021) and MacKinnon and Pirlott (2014) propose using treatment as an instrument for the mediator, which only works if treatment has no direct effect on the outcome and therefore does not apply to our setting. To the best of our knowledge, no causal mediation paper has proposed instruments that affect the mediator separately from treatment as a solution to separating direct and indirect effects.⁵

2 Bias Arising from Price Response

We consider a setting where J products are sold in T markets. All variables may vary by market but we omit the market subscripts to simplify notation. We study the impact of a marketing variable (treatment) $Treat_j$ on the demand for product j . The canonical example is advertising; however, our arguments apply for other marketing mix variables shifting demand that are likely to pass-through to prices, e.g., service effort, product quality, or product labeling.

The demand curve is given by

$$q_j = \alpha_0 - \alpha_p p_j + \alpha_T Treat_j + \xi_j, \quad (1)$$

where q_j and p_j denote quantity and price, respectively. The term ξ_j denotes an unobserved demand shock. We consider the case of downward sloping demand and a positive treatment effect, that is, $\alpha_p > 0$ and $\alpha_T > 0$. Throughout the paper we assume that treatment is exogenous, i.e., orthogonal to ξ_j . For instance, $Treat_j$ may be randomly assigned.

This general structure of demand encompasses two standard specifications: (i) a reduced-form linear demand equation (e.g. Porter (1983), DellaVigna and Gentzkow, 2019) with log-transforms of quantity and price; and (ii) a structural logit model, in which case the quantity variable is equal to $\ln(s_j) - \ln(s_0)$, as in Berry (1994).

The reduced-form pricing equation is given by

⁵The closest approach to ours is the “double randomization” approach (see e.g. Bullock et al. (2010)) where the mediator is randomly assigned in a separate experiment that is used only to measure the effect of the mediator on the outcome.

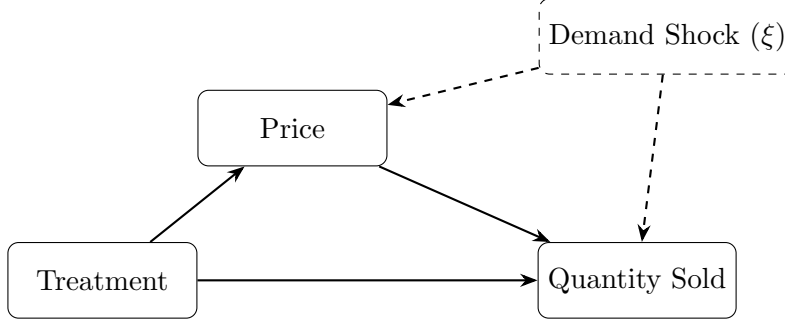


Figure 1: **Directed Acyclic Graph: Relationship between Treatment, Price, Quantity, and Unobserved Demand Shock.**

$$p_j = \beta_0 + \beta_T \text{Treat}_j + \beta_\xi \xi_j + \eta_j, \quad (2)$$

where η_j are supply shifters, such as cost shocks. The unobserved terms in the demand and supply equation, denoted by ξ_j and η_j , are assumed to be uncorrelated.

We expect that in most settings $\beta_\xi > 0$ because an increase in demand typically leads to higher prices. The coefficient β_T is likely to be non-zero for two reasons. First, treatment can affect prices indirectly via demand. For example, when treatment increases demand, it is plausible to expect a price increase, in which case β_T would be positive. Second, treatment can affect prices via the supply side. For instance, in our empirical application, we analyze the impact of supermarket feature advertising, which tends to be associated with lower prices. Such a price decrease likely occurs because manufacturers offer wholesale price discounts when their product is advertised. To accommodate both cases, we keep the derivations in this section general, and allow for positive and negative β_T .

Figure 1 represents an alternative way to illustrate the equations outlined above using a directed acyclic graph (DAG). The DAG depicts two causal pathways through which treatment affects quantity sold, namely, the direct effect and the indirect effect via prices. The direct effect of treatment on demand is depicted at the bottom of the graph and corresponds to α_T – the treatment coefficient in the demand equation. The indirect effect depends on the relationship between treatment and price (β_T) and the causal effect of price on demand ($-\alpha_p$). Importantly, price reacts to ξ , which also enters the demand equation. There are therefore three paths that connect treatment and quantity sold in this DAG: the direct causal path ($\text{Treatment} \rightarrow \text{Quantity}$), the indirect causal path ($\text{Treatment} \rightarrow \text{Price} \rightarrow \text{Quantity}$), and the backdoor path via price and the unobserved demand shock ($\text{Treatment} \rightarrow \text{Price} \leftarrow \xi \rightarrow \text{Quantity}$). We will refer back to the relationships depicted in the DAG throughout the paper.

2.1 Bias Without Price Control

First, we consider the case where the researcher regresses quantity on treatment without controlling for price. Such a regression estimates the relationship between treatment and demand after prices have adjusted in response to the treatment. We derive this relationship by substituting the pricing equation into the demand equation:

$$q_j = (\alpha_0 - \alpha_p \beta_0) + (\alpha_T - \alpha_p \beta_T) Treat_j + [(1 - \alpha_p \beta_\xi) \xi_j - \alpha_p \eta_j]. \quad (3)$$

If $\beta_T > 0$, then the impact of treatment is underestimated. In this case, treatment shifts the demand curve up, which results in an increase in prices (because $\beta_T > 0$), which in turn leads to a decrease in realized demand (because $-\alpha_p \beta_j < 0$). Hence the possible increase in demand due to treatment is “dampened” by the price increase. Conversely, if $\beta_T < 0$, as in the case of feature advertising, the treatment coefficient is overestimated because the positive demand effect of the price decrease is falsely attributed to treatment.

In terms of the DAG presented in Figure 1, this regression recovers a combination of the direct effect of treatment (α_T) and the indirect effect due to price adjustments ($-\alpha_p \beta_T$). In DAG terminology, price is a collider, meaning that the indirect path from advertising to quantity via the unobserved demand shock is blocked. As a result, only the direct and indirect causal paths are open.

We note that if the goal is to measure the impact of treatment on equilibrium quantity, then a simple regression of quantity on treatment will provide an unbiased estimate of this object. When we use the term “bias” with regards to a regression without price control, we refer to the case where the aim is to isolate the direct effect on demand.

2.2 Bias when Controlling for Price

Next, we turn to the scenario where the researcher regresses quantity on treatment while also controlling for price, which is endogenous and correlated with both treatment and the demand shock ξ_j (see equation (2) and Figure (1)).

To derive the bias that results from including the endogenous price control, we rely on standard properties of linear regression. In particular, one way to interpret the multi-variate regression of quantity on treatment and price is a two-step process: first, regress treatment on price, and then regress quantity on the residual from the first-stage regression. By the Frisch-Waugh Theorem (see [Frisch and Waugh, 1933](#)), the coefficient on the treatment variable in the second regression is the same as the multi-variate regression coefficient.

We let \widetilde{Treat}_j denote the residual from the first-stage regression:

$$\widetilde{Treat}_j = Treat_j - (\gamma_0 + \gamma_p p_j),$$

where γ_p will have the same sign as β_T , the treatment coefficient in the pricing equation.

The second-state coefficient of a regression of quantity on the first-stage residuals is given by:

$$\begin{aligned}\hat{\alpha}_T &= \frac{\text{Cov}(\widetilde{Treat}, q)}{\text{Var}(\widetilde{Treat})} = \frac{\text{Cov}(\widetilde{Treat}, \alpha_T \widetilde{Treat} + \xi)}{\text{Var}(\widetilde{Treat})} = \frac{\text{Cov}(\widetilde{Treat}, \alpha_T \widetilde{Treat} + \xi)}{\text{Var}(\widetilde{Treat})} \\ &= \alpha_T - \gamma_p \frac{\text{Cov}(p, \xi)}{\text{Var}(\widetilde{Treat})}\end{aligned}\tag{4}$$

$$= \alpha_T - \beta_T \beta_\xi \frac{\text{Var}(\xi)}{\text{Var}(\tilde{p})}\tag{5}$$

where the second equality in the first line follows from the fact that \widetilde{Treat} is uncorrelated with price. The third equality in the first line follows from the fact that $Treat$ can be decomposed into \widetilde{Treat} and the fitted value of the first stage regression ($\gamma_0 + \gamma_p p_j$) and again the fact that \widetilde{Treat} is uncorrelated with price. The expression in the second line follows from the fact that $Treat$ is uncorrelated with ξ (due to random assignment). The final line can be derived by plugging equation (2) into the covariance expression.⁶ The term \tilde{p} denotes the residual from a regression of price on treatment.

In the case of $\beta_T > 0$, it follows that $\hat{\alpha}_T < \alpha_T$, and the treatment effect is underestimated when including an endogenous price control in the regression. Conversely, if $\beta_T < 0$, it follows that $\hat{\alpha}_T > \alpha_T$, and the treatment effect is overestimated when including an endogenous price control in the regression. Therefore, the estimate of the treatment effect when including price as an endogenous control variable leads to a bias in the same direction as the bias from a regression without price control. Whether the bias is larger with or without price control is not immediately obvious when comparing equations (3) and (5). However, a re-arranged bias expression, which we present below, will allow us to compare the bias magnitude in both cases.

The presence of a bias in this regression is also visible in the DAG. As mentioned earlier, price is a collider because it depends on both the treatment and the demand shock. Conditioning on price has two effects. First, it closes the indirect causal path via price. However, at the same time, it opens the backdoor path via the demand shock ξ . The fundamental issue is therefore that either choice, omitting or including price leads to open paths causing bias. We either leave open the indirect causal path in the former case, or the backdoor path, previously blocked by price, in the latter case. The open paths in both versions of the regression, without price control and with an endogenous price control, correspond to the different bias formulas we derived above.

The expression in equation (5) provides insights into the conditions under which this bias occurs. In particular, the bias occurs only if price is correlated with both the treatment ($\beta_T \neq 0$) and the demand-side error ($\beta_\xi \neq 0$). Hence, the bias stems from the fact that two demand shifters, treatment and ξ , affect price due to the equilibrium response to changes in demand. In terms of the DAG representation, bias occurs only if there are arrows between price and treatment, and between price and the demand shock ξ . If either arrow was absent, price does not function as a

⁶We also write γ_p as a function of β_T and plug it into the relationship. See Appendix A for more details.

collider, in which case the regression with price control would produce an unbiased estimate of the direct causal effect of treatment.

Finally, to assess the relative magnitude of the bias we provide an alternative formulation where the bias in the treatment coefficient is expressed as a function of the bias in the price coefficient:⁷

$$\hat{\alpha}_T - \alpha_T = (\hat{\alpha}_p - \alpha_p)\beta_T = -\alpha_p\beta_T + \hat{\alpha}_p\beta_T \quad (6)$$

The first equality links the bias in the estimate of the treatment effect to the bias in the estimate of the price coefficient. The second equality highlights the relative magnitude of the bias in both regressions because the bias when not controlling for price is equal to $-\alpha_p\beta_T$.

As explained earlier, if price is positively related to the treatment, i.e., $\beta_T > 0$, the biases in the price coefficient and treatment effect in the regression controlling for the price have the same direction. That is, if the regression underestimates the price coefficient, it also underestimates the treatment effect. This is the most common case, as in most setups price endogeneity leads to attenuation of the price coefficient.⁸ In addition to the direction of the bias, it is useful to assess the relative magnitude of the bias in both regressions, which is equal to $\hat{\alpha}_p\beta_T$. If the estimate of price coefficient is negative, i.e., if $\hat{\alpha}_p < 0$, then the regression controlling for price has smaller bias. This is because the bias difference term is positive, and it mitigates the negative bias from the regression without the price. Conversely, if the price endogeneity is large enough to result in an upward sloping demand curve estimate, i.e., $\hat{\alpha}_p > 0$, a regression controlling for price produces more bias. In the knife edge case of $\hat{\alpha}_p = 0$, the estimates are the same in both regressions, which is intuitive because a zero price coefficient is equivalent to excluding price from the regression. Importantly, the same degree of the bias in both regression does not necessarily mean no bias; in other words, both regressions are likely to produce biased treatment effects, even if the estimated price coefficient is zero. Therefore, testing if the price coefficient in the demand regression is zero is not a valid test for the presence of the bias in the treatment effect.

The case of negative correlation between the price and treatment effect, that is if $\beta_T < 0$, is potentially more problematic. In this circumstance, the bias in treatment effect has the opposite sign compared to the bias in the price coefficient. For instance, the common case of an underestimated price coefficient results in an overestimation of the treatment effect, which may lead to false positives. In other words, we may detect positive treatment effects, when the true effect is absent, or even negative. This occurs in both regressions, as the bias in the regression without the price is given by $-\alpha_p\beta_T$. Again, the relative magnitude of the bias depends on the sign of the estimated price coefficient in the exact same way as for the case of $\beta_T > 0$. For example, consider a common situation, when the estimate of the demand curve is attenuated, but is still downward sloping, $\hat{\alpha}_p > 0$. We obtain a negative difference term $\hat{\alpha}_p\beta_T < 0$, and since the bias in the regression

⁷See Appendix A for a detailed derivation.

⁸In a less common case, when price coefficient is overestimated (which may occur under countercyclical pricing, or $\beta_\xi < 0$), the treatment effect is also biased upwards.

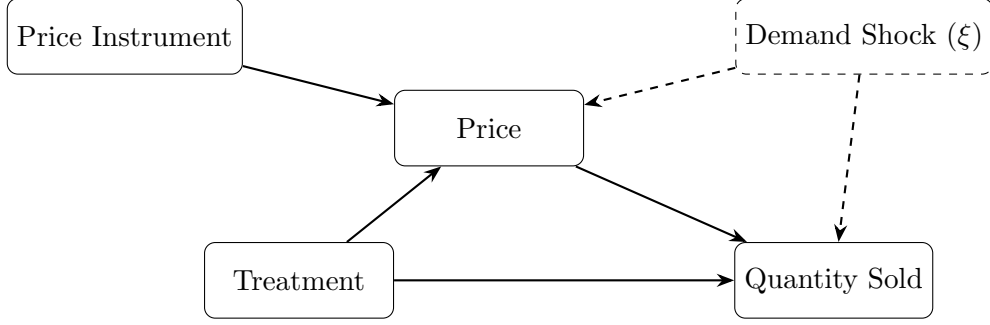


Figure 2: **Amended DAG with Price Instrument.**

without price is positive, the corresponding bias in a regression with price would be smaller.⁹

In summary, our derivations provide testable guidelines to the researcher about the direction and relative degree of the bias when including and omitting price in the regression. Primarily, the direction of the bias depends on the correlation of prices and treatment, which is observed in the data. Further, the relative bias magnitudes, when including and omitting price, depend on the sign of the estimate of price coefficient, which is also observable. Specifically, if the regression recovers downward sloping demand, it would typically lead to lower bias in the treatment effect. Perhaps more importantly, these relationships also deliver a formal test for the bias, which we present in Section 2.4.

2.3 Instrumenting Price

Equation (4) points towards a possible way to obtain an unbiased estimate of the direct effect of treatment on demand. The bias in this equation arises because price is correlated with the unobserved demand shock ξ . Therefore, standard price instruments, i.e., variables that are correlated with price but not with the demand shock, allow the researcher to obtain an unbiased estimate of the treatment effect. When instrumenting price, the expression in equation (4) will depend on the covariance of *instrumented* price with the demand shock. If the instrument is valid, the covariance will be zero and therefore the entire bias term is equal to zero. An alternative way to obtain the same conclusion is to examine equation (6), which shows that the bias in the treatment coefficient disappears when the price coefficient is unbiased. An unbiased estimate of price coefficient can be obtained by employing a valid price instrument.

Further, inspecting the DAG relationships provides evidence for the empirical strategy of using a price instrument. If the instrument is uncorrelated with the demand shock, as depicted in Figure 2, then it isolates the variation in price that is uncorrelated with the demand shock. This feature is indicated in the graph by the absence of an arrow from the demand shock to the instrument. The instrument thus allows us to control for price, by blocking the indirect causal pathway, without opening the backdoor path via the demand shock.

⁹Again, if the recovered demand curve is upward sloping, the estimate without the price controls is preferable.

2.4 Testing for Endogenous Price Response

A simple way to test for the possible bias in the treatment effect without accounting for an endogenous price response is to regress price on treatment. This regression recovers an unbiased estimate of β_T due to the random assignment of treatment. If $\beta_T = 0$, then no bias arises when regressing quantity on treatment, or when controlling for price without an instrument. Inspecting equations (3) and equation (6) shows that the bias terms disappear when $\beta_T = 0$ holds. In this case, a simple regression of quantity on treatment delivers an unbiased estimate of the direct treatment effect, as the indirect effect is zero.¹⁰

Two other conditions lead to an absence of bias: when the true price coefficient is zero, i.e., if $\alpha_p = 0$, no bias occurs in the regression without price control, and if $\beta_\xi = 0$ the regression with an uninstrumented price control is unbiased. Both cases are not directly testable, as the true price coefficient and demand shocks are typically unobserved. Both cases are also unlikely to apply in most settings, because they imply either that demand does not react to price, or that prices do not react to demand shocks.

Finally, we note that finding a similar value of the treatment coefficient in a regression without price and regression with an endogenous price control is not sufficient to conclude that the treatment coefficient is unbiased. Because both coefficients are biased, it is possible that the treatment coefficient is equally biased in both cases. As we showed earlier (see equation (6)), this case arises when the price coefficient is equal to zero.

3 Simulations

In this section, we illustrate the size and direction of the possible bias in specifications with and without a price control using a set of numerical simulations. We start with simulations based on linear demand and supply equations that match the ones in the previous section and then extend the analysis to a structural price setting model and a more flexible demand model with heterogeneous consumers.

3.1 Reduced-form Pricing Equation

We consider a standard micro-foundation for the demand function in equation (1). Suppose that each consumer has a utility for product j given by

$$u_{ijt} = \alpha_0 - \alpha_p p_{jt} + \alpha_T Treat_{jt} + \xi_{jt} + \varepsilon_{ijt}, \quad (7)$$

where ε_{ijt} is distributed as type-1 extreme value. Consumers can also choose an outside option denoted by $j = 0$, with normalization $u_{i0t} = \varepsilon_{i0t}$. Consumers purchase one product per period

¹⁰As we show later, there still may be an advantage of including price as a control in this setting, even if the price coefficient is biased due to endogeneity. This is because, including the price may lower the variance of the error term, and increase power when estimating the treatment effect.

that yields the highest utility. Denote market share of the product j in market t as s_{jt} . Berry (1994) shows that in this case the aggregate demand function can be linearized. We supplement the demand relationship with a reduced-form pricing equations.

The parametrization of demand and supply is given by:

$$\begin{aligned}\ln(s_{jt}) - \ln(s_{0t}) &= \alpha_0 - \alpha_p p_{jt} + \alpha_T \textit{Treat}_{jt} + \xi_{jt} \\ p_{jt} &= \beta_0 + \beta_T \textit{Treat}_{jt} + \beta_\xi \xi_{jt} + \beta_Z Z_{jt} + \eta_{jt},\end{aligned}\tag{8}$$

which is identical to the framework outlined in the previous section, except for the fact that we use the difference in log-shares as the outcome variable in the demand equation. This transformation allows us to obtain a structural interpretation of the regression coefficients as utility parameters. We also introduce an instrument in the supply equation denoted by Z_{jt} . The instrument is a cost shifter uncorrelated with ξ and η . We use the instrument to demonstrate how to obtain an unbiased estimate of α_T .

We simulate treatment, prices and consumer choices for 1,000 markets with two inside goods. We set the following default values of the parameters on the demand side: $\{\alpha_0 = 10, \alpha_p = 1, \alpha_T = 1, \sigma_\xi = 1.2\}$, and on the supply-side: $\{\beta_0 = 10, \beta_T = 0.5, \beta_\xi = 0.4, \sigma_\eta = 1\}$ in all simulations, unless specified otherwise.¹¹ We assume that treatment is a dummy variable that takes a value of 1 with 50% probability and that it is IID across firms and markets. We consider a set of different cases that generate different biases (or an absence of bias) depending on the values of β_T , β_ξ and σ_ξ .

Unbiased Estimates Without Instruments We start by considering two special cases in which it is not necessary to instrument for price. In the first case, we simulate data for a model where pricing does not react to treatment, i.e. $\beta_T = 0$. We estimate the model without controlling for price and controlling for price without an instrument and report results in columns (2) and (3) of Table 1. Given our assumption that price and treatment are uncorrelated in the data-generating process, it is unsurprising that the estimated treatment effect is close to the true value in both regressions. The price coefficient in column (3) is biased due to the correlation of price the demand shock, which introduces the standard price endogeneity concerns. Crucially, this endogeneity does not influence the estimation of the treatment coefficient, as treatment remains uncorrelated with price. This follows directly from equations (3) and (4), where the bias terms disappears if $\beta_T = 0$. Including price in the regression enhances the precision of the (unbiased) treatment effect estimate by reducing its standard error, even though the price coefficient itself is biased. We also present results for the case in which price is instrumented, despite instrumenting being unnecessary in this context. As expected, the treatment coefficient remains unbiased; however, its standard error increases relative to the regression that includes price as an uninstrumented control due to decrease in power caused by the (unnecessary) first-stage regression.

¹¹ ξ_{jt} and η_{jt} are drawn from normal distributions with mean zero and the indicated variance. Z_{jt} is drawn from a standard normal distribution.

| | | (1) | (2) | (3) | (4) |
|--|-----------|-------------|------------------|--------------------------|-------------------------------|
| | | True Values | No Price Control | Endogenous Price Control | Price Control with Instrument |
| (a) No Correlation between Price and Treatment $\beta_T = 0$ | Treatment | 1 | 1 (0.057) | 1 (0.049) | 1 (0.054) |
| | Price | -1 | | -0.56 (0.021) | -1 (0.088) |
| (b) No Price Endogeneity $\beta_\xi = 0$ | Treatment | 1 | 0.5 (0.070) | 1 (0.056) | 1 (0.071) |
| | Price | -1 | | -1 (0.026) | -1 (0.089) |
| (c) Low Price Endogeneity $\sigma_\xi = 1.2$ | Treatment | 1 | 0.5 (0.057) | 0.78 (0.050) | 1 (0.070) |
| | Price | -1 | | -0.56 (0.022) | -1 (0.091) |
| (d) Medium Price Endogeneity $\sigma_\xi = 2.14$ | Treatment | 1 | 0.5 (0.076) | 0.5 (0.077) | 1 (0.127) |
| | Price | -1 | | 0 (0.027) | -1.01 (0.162) |
| (e) Large Price Endogeneity $\sigma_\xi = 2.7$ | Treatment | 1 | 0.5 (0.085) | 0.35 (0.086) | 1.01 (0.158) |
| | Price | -1 | | 0.29 (0.028) | -1.02 (0.216) |
| (f) Negative Correlation between Price and Treatment $\beta_T = -0.5$ | Treatment | 1 | 1.5 (0.054) | 1.22 (0.049) | 1 (0.072) |
| | Price | -1 | | -0.56 (0.022) | -1 (0.093) |

Table 1: **Estimation Results Based on Simulated Data.** Parameters taken on the following default values on the demand side: $\{\alpha_0 = 10, \alpha_p = -1, \alpha_T = 1, \sigma_\xi = 1.2\}$ and on the supply-side: $\{\beta_0 = 10, \beta_T = 0.5, \beta_\xi = 0.4, \sigma_\eta = 1\}$ in all simulations. Any deviations from default parameters are indicated in the left-most column.

In the second case, we set $\beta_\xi = 0$, which implies that pricing reacts to treatment but not to demand shocks. While this may appear like an unusual assumption, we believe it serves as a useful benchmark for the reasons we outline below. In this setting, a regression without the price control leads to an underestimation of the treatment coefficient because price reacts positively to treatment. This is in contrast to the first case, where treatment did not influence price, allowing for unbiased estimation even without price control. When controlling for price, see column (3), we obtain an unbiased estimate of the treatment effect. Notably there is no need to instrument for price because price does not react to demand shocks.¹² This scenario is particularly noteworthy

¹²Instrumenting for price also leads to an unbiased treatment coefficient, but increases the standard error of the treatment coefficient because it utilizes only part of the variation in the (exogenous) price control variable.

because it represents the only special case in which including a price control is necessary and where the inclusion of the uninstrumented price control yields an unbiased treatment coefficient. Although it is common in empirical work to control for price without instrumenting, we believe the conditions under which this specification is valid are restrictive and unlikely in most real-world settings. Specifically, for this approach to be correct, price must respond exclusively to treatment and remain unaffected by any other demand shocks.

Downward Bias: Positive Correlation Between Price and Treatment Next, we simulate data using different values of $\sigma_\xi = \{1.2, 2.14, 2.7\}$, while setting β_T and β_ξ to their default values. The results are presented in panels (c) to (e) of Table 1. Intuitively, regulating the variance of the unobserved demand shock, regulates the severity of price endogeneity. As the variance of ξ increases, a larger part of the total price variation is driven by the unobserved shock, thus leading to a larger bias in the price coefficient, holding everything else constant.

For a low value of σ_ξ , we find that the bias in the treatment coefficient is smaller with an endogenous price control than in a regression without a price control. As discussed earlier, equation (6) reveals that the bias in the treatment coefficient is related to the bias in the price coefficient. Our numerical results confirm that the underestimation of the price coefficient leads to an underestimation of the treatment coefficient. Further, despite the bias, the price coefficient remains negative. Thus, according to our earlier result, including the endogenous price control shrinks the bias in the treatment effect estimate. Instrumenting for price, however, fully resolves the endogeneity issue and produces an unbiased estimate of the treatment effect.

In panel (d), we increase the variance of the demand shock which leads to a larger downward bias in the price coefficient because the influence of the demand shock on price grows. We intentionally set the variance such that the estimated price coefficient is approximately zero. This case is logically equivalent to omitting price from the regression altogether, implying that the bias in the treatment coefficient is the same in columns (2) and (3). Importantly, even in this case, the bias remains large. As before, instrumenting leads to unbiased estimates of both the price and treatment coefficient.

Finally, in panel (e) we induce high enough variance that so that the estimated price coefficient becomes positive. In this scenario, the treatment coefficient is relatively more biased when using an endogenous price control versus a regression without price. As we discussed earlier, this case is detectable because a positive price coefficient should immediately signal that the price coefficient is biased.

In summary, the three cases in panels (c) to (e) highlight that if the price coefficient estimate is negative, the relative bias when controlling for prices without instrumenting tends to be lower than when not controlling for price. The case of zero price coefficient is particularly noteworthy, because a researcher running regressions with and without price control might be tempted to conclude that the price “does not matter.” Further, the researcher may deduct that the treatment coefficient is unbiased, because it is “robust” to the inclusion of the price control. However, this interpretation is incorrect, as both specifications are equally biased. Finally, when the sign of the price coefficient

becomes positive, controlling for prices without an instrument exacerbates the bias in the treatment coefficient. This final case is easy to detect because a positive price coefficient should signal an obvious endogeneity problem.

Upward Bias: Negative Correlation Between Price and Treatment In panel (f) of Table 1, we consider the case where price and treatment are negatively correlated. This case behaves similarly to the positive correlation case, but with the reverse direction of the bias. We set $\beta_T = -0.5$, and $\sigma_\xi = 1.2$, making this dataset equivalent to the one in panel (c), except that price and treatment are now negatively correlated. When regressing log-shares on treatment alone, the treatment coefficient is biased upwards. This upward bias occurs because treatment is accompanied by lower prices, and hence, both treatment and the price reduction increase market shares. When controlling for price without an instrument, the upward bias is reduced but not fully eliminated. As in all other cases, using an instrument for price leads to unbiased estimates of the treatment and price coefficients. Although not reported in the table, setting the variance of the demand shock high enough can change the sign of the estimated price coefficient. When this occurs, controlling for endogenous prices worsen bias instead of reducing it. This case is identical to the case in panel (e) but with a bias occurring in the opposite direction.

3.2 Bertrand-Nash Competition

In this section, we examine a micro-founded pricing model based on a price-setting game, following [Berry et al. \(1995\)](#) (henceforth, BLP).

Logit Demand We start by considering the same demand specification as in the previous section. That is, the utility is given by equation (7), which generates standard logit market shares. Further, we assume that firms face marginal cost, represented by

$$c_{jt} = \bar{c} + \gamma_T \text{Treat}_{jt} + \gamma_Z Z_{jt} + \eta_{jt}, \quad (9)$$

where Z_{jt} are cost shifters observed by the econometrician, assumed to follow a standard normal distribution, while η_{jt} denotes unobserved cost shifters, which are log-normally distributed (obtained by exponentiating a normal distribution with standard deviation of 0.3). Notably, Z_{jt} serves a role analogous to the reduced-form price shifters in equation (8); however, unlike the reduced-form price shifters, these cost shifters are structural, as they have an economic interpretation in terms of marginal cost. We set $\{\bar{c} = 10, \gamma_T = 0, \gamma_Z = 0.3\}$ as the default parameters. The value of γ_Z and the distribution of η_{jt} remain unchanged throughout the paper.

We derive equilibrium prices by solving the first-order conditions for profit maximization, given by the system of equations

$$\frac{\partial s_{jt}(p_{jt}, p_{-jt}; \text{Treat}_{jt})}{\partial p_{jt}} (p_{jt} - c_{jt}) + s_{jt} = 0, \quad (10)$$

where p_{-jt} is the price charged by the competitor. We assume that firms take competitors' prices as given, implying that the solution to the above equations (subject to second-order conditions) represents the Bertrand equilibrium of the pricing game.¹³

This specification allows for two pathways by which treatment affects prices. First, treatment affects market shares via the utility function, altering the demand elasticity, and thereby affecting the solution to the first-order conditions – a mechanism we refer to as the demand effect. Second, treatment may directly affect marginal costs, as modulated by the parameter γ_T . The resulting change in marginal costs is subsequently passed through to prices, which we term the supply effect. The latter mechanism is relevant in contexts, such as promotions, when manufacturers can offer retailers a discounted wholesale price when a product is advertised or featured (see [Blattberg and Levin, 1987](#); [Kumar et al., 2001](#)). Both the demand and supply effects provide micro-foundations for the parameter β_T in equation (8), generating either a positive or negative correlation between treatment and price. Finally, we note that the first-order conditions dictate the dependence of prices on ξ , which provides micro-foundations for the parameter β_ξ .

The structural parameters are calibrated to align with the various cases in the reduced-form simulations presented in Table 1, with the exception of the first case, in which price and treatment are uncorrelated.¹⁴ Consistently with the reduced-form analysis, we manipulate the degree of price endogeneity by adjusting the variance of the demand shifter ξ . The first set of simulations excludes supply-side effects of treatment by setting $\gamma_T = 0$.

The first row of Table 2 examines a scenario where $\sigma_\xi = 0$, eliminating price endogeneity. A key distinction from the reduced-form approach is worth noting: in the reduced-form framework, the absence of endogeneity was achieved by setting $\beta_\xi = 0$, which allows for variation in ξ while mechanically shutting down an endogenous price response. However, replicating this setup within a structural supply-side pricing model is challenging, as it would require assuming that firms do not incorporate ξ into their pricing decision – a deviation from the canonical supply model. Instead, we eliminate price endogeneity by setting the variance of ξ to zero. While this setting is somewhat artificial and possibly less practically relevant, we include it as a benchmark for the comparison with the subsequent results. The simulation under this configuration yields a positive correlation between prices and treatment. This outcome occurs because treatment leads to an outward shift in demand, reducing the demand elasticity for any fixed price.¹⁵ One can show that less elastic demand leads to a higher price best response, and as a result, equilibrium prices for both companies increase. The results closely mirror those obtained using the reduced-form analysis. Both the OLS regression with a price control and the IV estimator produce unbiased estimates of the treatment effect.¹⁶

¹³The system of equations is solved using a standard Newton method. We verify second-order conditions for every solution.

¹⁴This case is omitted due to the trivial conclusions when price and treatment are uncorrelated. We note that in the setting with a structural supply model, such scenario would only arise if the supply and demand effects of treatment on prices cancel each other.

¹⁵In the logit model, price elasticity is given by $-\alpha p p_{jt}(1 - s_{jt})$. Since the treatment increases s_{jt} for any given price, the demand becomes less elastic.

¹⁶Notably, in the specifications with price control (OLS and IV), the standard errors are zero, as we set the variance

| | | (1) | (2) | (3) | (4) |
|--|--------------|-------------|------------------|--------------------------|-------------------------------|
| | Coefficients | True Values | No Price Control | Endogenous Price Control | Price Control with Instrument |
| (b) No price endogeneity $\sigma_\xi = 0$ | Treatment | 1 | 0.65 (0.008) | 1 (0.00) | 1 (0.00) |
| | Price | -1 | | -1 (0.00) | -1 (0.00) |
| (c) Low price endogeneity $\sigma_\xi = 0.6$ | Treatment | 1 | 0.65 (0.011) | 0.82 (0.013) | 1 (0.017) |
| | Price | -1 | | -0.47 (0.014) | -1 (0.034) |
| (d) Medium price endogeneity $\sigma_\xi = 1.4$ | Treatment | 1 | 0.63 (0.023) | 0.63 (0.025) | 1.01 (0.045) |
| | Price | -1 | | 0 (0.011) | -1 (0.088) |
| (e) Large price endogeneity $\sigma_\xi = 2.9$ | Treatment | 1 | 0.63 (0.041) | 0.58 (0.043) | 1.00 (0.082) |
| | Price | -1 | | 0.13 (0.026) | -0.99 (0.160) |
| (f) Negative correlation between price and treatment $\sigma_\xi = 1.5, \bar{c} = 3, \gamma_T = -2$ | Treatment | 1 | 1.89 (0.026) | 1.27 (0.012) | 1 (0.084) |
| | Price | -1 | | -0.27 (0.012) | -1 (0.089) |

Table 2: **Simulations with Structural Supply & Logit Demand Model.** Parameters taken on the following default values on the demand side: $\{\alpha_0 = 10, \alpha_p = -1, \alpha_T = 1\}$ and on the supply-side: $\{\bar{c} = 1, \gamma_T = 0, \gamma_Z = 0.3\}$ in all simulations except when different values are indicated in the left-most column. We label the cases to mirror the reduced-form results in Table 1 and therefore start with case (b).

To explore varying degrees of price endogeneity, we adjust the variance of unobserved demand heterogeneity by setting $\sigma_\xi = \{0.6, 1.4, 2.9\}$. As in the reduced-form case, each of these scenarios produces a positive correlation between prices and treatment. The bias patterns replicate the corresponding reduced-form results in Table 1. Specifically, models excluding price, as well as OLS models incorporating price, tend to underestimate the treatment effect. When price endogeneity becomes more pronounced, the inclusion of price exacerbates bias, but only when the estimated price coefficient is positive. Finally, we observe substantial bias even in cases where the price coefficient in the OLS model is zero (see case (d)), again mirroring the equivalent result from the reduced-form simulations.

We also replicate a scenario in which price is negatively correlated with treatment, as in the last of ξ to zero. This is because removing ξ leads to deterministic market shares; thus, the Berry logit inversion produces an OLS regression without residuals.

row of Table 1. This case cannot be achieved within the standard logit model without supply-side effects, as the demand effect alone consistently induces a positive correlation between treatment and price. There are two potential approaches to address this challenge. The first involves enriching the demand model to allow for a rotation of the demand curve induced by treatment. Such a rotation could increase the demand elasticity for the marginal consumer, thereby generating a negative correlation between optimal prices and treatment. The second approach introduces supply-side effects, such as trade promotions. Since supply-side effects preserve the linear demand assumption, they provide a natural starting point, maintaining linear demand and ensuring compatibility with the theoretical framework outlined in Section 2 and the reduced-form analysis in Section 3.1.

In the last row of Table 2, we consider a case where $\bar{c} = 3$ and $\gamma_T = -2$, which produces a negative correlation between prices and treatment. In line with the reduced-form results, we replicate the upward bias observed in the treatment effect estimate, both in the regression without the price and in an OLS regression with an endogenous price control.

Heterogeneous Preferences Next, we consider a richer demand model with random coefficients, akin to the BLP framework. This extension is motivated by two objectives. First, it allows us to demonstrate that the bias of non-IV estimators persists even in a non-linear demand setting, when the theoretical results from Section 2.2 are not directly applicable. Second, it enables us to consider the case in which prices and treatment are negatively correlated without relying on supply-side effects. As noted earlier, achieving a negative correlation requires a rotation of the demand curve due to treatment, which may arise if the price coefficient varies heterogeneously across consumers.

To this end, we consider the simplest version of a random coefficient logit model with two latent consumer classes (henceforth referred to as segments), differentiated by their intercepts and price coefficients. To maintain focus on the estimation bias of a single parameter, we exclude heterogeneous treatment coefficients, thereby facilitating a clearer discussion of bias direction. The utility function is specified as:

$$u_{ijt} = \alpha_{0g(i)} - \alpha_{pg(i)}p_{jt} + \alpha_T \text{Treat}_{jt} + \xi_{jt} + \varepsilon_{ijt},$$

where $g(i) \in 0, 1$ is the segment of consumer i . Market shares for each segment, s_{gjt} , are derived using the standard logistic formula. The aggregate market share is computed as a weighted sum of the segment-specific shares: $s_{jt} = \sum_g s_{gjt}P_g$, where P_g denotes the proportion of consumers in segment g . For simplicity, we assume $P_g = 0.5$, and that this value is known to the econometrician. This assumption ensures a conservative assessment of estimation bias by eliminating potential misspecification due to errors in segment size estimation. Additionally, we assume that the econometrician observes only aggregate market shares, not segment-specific shares. The marginal cost specification follows equation (9), with parameter values set to $\{\bar{c} = 1, \gamma_T = 0, \gamma_Z = 0.3\}$.

We classify consumers into two segments: “loyals,” and “switchers.” . Loyals are an inelastic group, characterized by a high baseline utility for inside goods (or a low utility for the outside option), while switchers are an elastic group, with low baseline utility for inside goods (or a high

| | Coefficients | True Values | No Price Control | Endogenous Price Control | Price Control with Instrument |
|--|--------------|-------------|------------------|--------------------------|-------------------------------|
| BLP with positive correlation between price and treatment $\rho(p, \text{Treat}) = 0.28$ $\rho(p, \xi) = 0.25$ | Treatment | 2 | 1.21 (0.019) | 1.75 (0.028) | 1.97 (0.035) |
| | Intercept | 6 | -0.94 (0.145) | 2.65 (0.298) | 5.51 (0.203) |
| | - Loyals | | | | |
| | Intercept | 3 | -2.08 (0.160) | 4.46 (0.4151) | 3.10 (0.188) |
| | - Switchers | | | | |
| | Price | -2.3 | | -1.15 (0.103) | -2.12 (0.068) |
| | - Loyals | | | | |
| | Price | -2.3 | | -2.72 (0.158) | -2.32 (0.083) |
| | - Switchers | | | | |
| BLP with negative correlation between price and treatment $\rho(p, \text{Treat}) = -0.19$ $\rho(p, \xi) = -0.09$ | Treatment | 2 | 1.51 (0.027) | 1.44 (0.018) | 1.93 (0.042) |
| | Intercept | 6 | 2.82 (0.067) | 2.64 (0.162) | 5.43 (0.201) |
| | - Loyals | | | | |
| | Intercept | 3 | -4.15 (0.036) | -1.02 (0.056) | 2.70 (0.191) |
| | - Switchers | | | | |
| | Price | -0.2 | | 0.067 (0.013) | -0.15 (0.016) |
| | - Loyals | | | | |
| | Price | -2.3 | | -0.82 (0.018) | -2.18 (0.042) |
| | - Switchers | | | | |

Table 3: **Simulations with Structural Supply & BLP Random Coefficients Demand Model.** Supply-side parameters taken on the following values: $\{\sigma_\xi = 0.8, \bar{c} = 1, \gamma_Z = 0.3, \gamma_T = 0\}$ in all simulations.

utility for the outside option). Although treatment affects the utility of both groups equally, its impact on market share is more pronounced for switchers due to the non-linear effect of utility on market shares in the logit framework. Consequently, when the treatment is active, the set of marginal consumers contains a relatively larger proportion of switchers. This leads to a rotation of the aggregate demand curve when treatment is active. Consequently, as more switchers become marginal, the firm faces downward pricing pressure. Simultaneously, the firm faces incentives to raise prices to capitalize on the outward shift in demand by exploiting the increased willingness to pay of the infra-marginal loyals. The net effect of treatment on the the price effect depends on the relative magnitudes of these two opposing forces, the demand rotation and the outward shift in demand.

We estimate the model using a Generalized Method of Moments (GMM) approach, operationalized through a nested fixed-point inversion of market shares. In the inner loop, we obtain a vector, ξ_{jt} , by inverting aggregate marker shares s_{jt} for a given set of parameters $\alpha = \{\alpha_{00}, \alpha_{01}, \alpha_{p0}, \alpha_{p1}, \alpha_T\}$. In the outer loop, we estimate the parameters by matching the moment

conditions $E[\xi_{jt}\mathbf{I}_{jt}; \boldsymbol{\alpha}] = 0$, where \mathbf{I}_{jt} is a vector of GMM instruments.¹⁷ We employ optimal instruments for prices and costs shifters to eliminate any degree of freedom in selecting the functional form of the instruments. To ensure robustness, we also repeated the analysis using BLP instruments and found numerically similar results. The choice of instruments varies depending on the specification of the model.

- Model without price controls: We use six instruments: a vector of ones, the firm’s own treatment Treat_{jt} , and the competitor’s treatment Treat_{-jt} , each interacted with firm identity (3×2 vectors).
- Model with endogenous price controls: We use a vector of ones, optimal instruments for price defined by $E\left[\frac{\partial \xi_{jt}}{\partial \alpha} | p_{jt}\right]$ and $E\left[\frac{\partial \xi_{jt}}{\partial \alpha} | p_{-jt}\right]$ (5-parameter vector per each conditional expectation), the firm’s own treatment Treat_{jt} , and the competitor’s treatment Treat_{-jt} , resulting in 13 moments.
- Model with exogenous cost shifters: We replace optimal price instruments that condition on price with instruments that condition on exogenous cost shifters. Formally, the new optimal instruments are given by $E\left[\frac{\partial \xi_{jt}}{\partial \alpha} | Z_{jt}\right]$ and $E\left[\frac{\partial \xi_{jt}}{\partial \alpha} | Z_{-jt}\right]$.

In the model with endogenous price controls, we deliberately use misspecified moment conditions by constructing the optimal instruments conditional on prices, p_{jt} . Since prices are endogenous, these optimal instruments are not orthogonal to the error term ξ_{jt} . This exercise mimics the orthogonality violation in the “un-instrumented” OLS case and is therefore expected to produce a biased estimate of the treatment effect.¹⁸

All estimators follow a two-step procedure to derive the optimal weighting matrix. In the first step, we use an initial weighting matrix (a diagonal matrix with inverses of moment condition variances, evaluated at the true parameters) to obtain preliminary parameter estimates. In the second step, we compute the optimal weighting matrix based on these estimates and re-estimate the model to improve efficiency. Standard errors are computed using a parametric bootstrap.

Table 3 presents two representative cases. In case 1, loyals and switchers are equally price-elastic. In this scenario, treatment leads to a price increase, and the estimated treatment effects are biased downward. This outcome replicates the reduced-form results. In case 2, loyals are highly inelastic, leading to a demand rotation as switchers enter the market. In this scenario, treatment induces price reductions to attract switchers. However, the treatment effects are again underestimated.

Notably, the bias direction in case 2 is opposite to previous examples (see case (f) in Tables 1 and 2), in which a negative correlation of price with treatment resulted in upwards bias. This dis-

¹⁷We use the term “GMM instruments” in a broad sense, referring to any variables that interact with the residuals to construct moment conditions. In this context, GMM instruments encompass not only traditional exogenous variables but also price instruments, denoted by Z , which serve to address endogeneity in price estimation.

¹⁸A typical estimation procedure with optimal instruments is to obtain a preliminary estimate of α using sub-optimal instruments, construct optimal instruments, and then re-estimate α . Our optimal instruments are evaluated at true value of α to avoid applying additional bias from mis-estimation of the derivative of ξ due to price endogeneity. The conditional expectations are estimated using separate kernel regressions employing squared exponential kernel with standard deviation of 0.1.

crepancy is likely driven by the countercyclical pricing property of this parameterization, wherein ξ and prices exhibit a negative correlation ($\rho = -0.09$). According to equation (4), the bias sign is determined by the correlation between prices and treatment, as well as the term $\text{cov}(\xi, p)$. In prior cases, the covariance term was always positive, but the demand rotation in the random coefficient model leads to a negative correlation. Crucially, since $\text{cov}(\xi, p)$ is unobserved, determining the bias direction in practice is challenging and requires additional assumptions about the unobserved relationship between prices and demand shocks.

4 Application: Impact of Feature Advertising on Demand

In this section, we apply our framework to a core marketing question: the estimation of the causal impact of feature advertising on demand for the featured product. We estimate demand for individual products using panel data at the store/week-level and a standard two-way fixed effect regression model:

$$q_{jt} = \alpha_{feat}Feat_{jt} + \alpha_{price}p_{jt} + \delta_j + \theta_t + \varepsilon_{jt}, \quad (11)$$

where δ_j and θ_t denote store and week fixed effects, respectively. The coefficient of primary interest is the feature coefficient, α_{feat} . We use log-transforms of purchase quantities and price. Following our earlier discussion, we show results for three specifications: (i) without price control, (ii) with an endogenous price control, and (iii) with instrumented price. We note that the specification in equation (11) is widely used in marketing research (e.g. Seiler and Yao (2017), Thomas (2020), Jindal et al. (2020), Goli et al. (2022)). Often, a similar model is embedded within a discrete-choice framework, where utility is specified as a function of price, advertising, and product fixed effects or product characteristics (see, e.g., Chintagunta and Dubé (2005), Chintagunta et al. (2005), Hendel and Nevo (2006)).

Consistent with the previous literature, we assume that the variation in feature advertising is as-good-as-random conditional on fixed effects. The key distinction in our analysis is the focus on the role of endogenous prices. Specifically, while we assume that feature advertising is uncorrelated with the error term, we allow for price to be potentially correlated with the regression error. This scenario is plausible because prices are easier to adjust in response to local demand shocks, and they vary at the store/week level. In contrast, feature advertising decisions are typically made at a higher geographic level for all stores belonging to the same chain, and are often predetermined as part of annual promotional calendars negotiated between manufacturers and retailers (Anderson et al. (2017)).¹⁹

We use retail scanner data from Nielsen-Kilts, which covers a broad set of retail stores in the US and focus our analysis on the top three products in seven product categories.²⁰ We start by

¹⁹If price endogeneity is absent and our instrument is valid, instrumenting for price should not alter the estimated coefficients on price and feature advertising. Thus, while we do not assume price endogeneity conditional on fixed effects, we allow for this possibility.

²⁰For each product we only retain stores where non-zero sales occur in at least 45 out of 52 weeks. We implement

| | (1) | (2) | (3) |
|---------------------------------|----------------------|----------------------|----------------------|
| Price/Feat. Correlation | | | |
| Dependent Variable | Log Price | Log Price | Log Price |
| Feature Dummy | -0.263*** (0.002) | -0.182*** (0.002) | -0.217*** (0.002) |
| No Price Control | | | |
| Dependent Variable | Log Quantity | Log Quantity | Log Quantity |
| Feature Dummy | 1.082*** (0.011) | 0.681*** (0.006) | 1.103*** (0.007) |
| Endogenous Price Control | | | |
| Dependent Variable | Log Quantity | Log Quantity | Log Quantity |
| Feature Dummy | 0.454*** (0.026) | 0.249*** (0.006) | 0.603*** (0.017) |
| Log Price | -2.386*** (0.097) | -2.376*** (0.026) | -2.302*** (0.076) |
| Store Fixed Effects | Yes | Yes | Yes |
| Week Fixed Effects | Yes | Yes | Yes |
| Observations | 198,160 | 129,690 | 261,526 |

Table 4: **Feature Ads: Price Correlation & Biased Sales Regressions.**

presenting results for the top three products in the Mayo category as an illustrative example before reporting results across all categories.

Before discussing the main regression results, we first estimate the correlation between features and price conditional on fixed effects by regressing price on the feature dummy, as well as store and week fixed effects. These results are reported in the top panel of Table 4. Each column represents one of the top three products in the category. For all three products, we find a negative and statistically significant relationship between feature advertising and price. A product being featured corresponds to approximately a 18%-26% price reduction, which is consistent with the typical size of temporary discounts that often accompany feature advertisements.

4.1 Biased Regression Specifications

Next, we proceed with the main regression analysis, and start by estimating equation (11) without including price in the regression. The results are reported in middle panel of Table 4. We find a large and statistically significant feature coefficient across all three products. We then add price

this selection step because price is not observed for weeks with zero sales. Only weeks with non-zero sales are included in our final sample.

| | (1) | (2) | (3) |
|---------------------|----------------------|----------------------|----------------------|
| 1st Stage | | | |
| Dependent Variable | Log Price | Log Price | Log Price |
| Feature Dummy | -0.012*** (0.001) | -0.067*** (0.003) | -0.047*** (0.002) |
| Cost Advantage | 0.973*** (0.004) | 0.834*** (0.007) | 0.896*** (0.005) |
| Partial F-Stat | 386,295 | 143,959 | 215,521 |
| 2nd Stage | | | |
| Dependent Variable | Log Quantity | Log Quantity | Log Quantity |
| Feature Dummy | 0.264*** (0.007) | 0.217*** (0.006) | 0.397*** (0.005) |
| Log Price | -3.107*** (0.021) | -2.552*** (0.031) | -3.251*** (0.019) |
| Store Fixed Effects | Yes | Yes | Yes |
| Week Fixed Effects | Yes | Yes | Yes |
| Observations | 198,160 | 129,690 | 261,526 |

Table 5: **Feature Ads: Instrumental Variable Regressions.**

to the regression without instrumenting, thus treating it as an endogenous control. The results are presented in the lower panel of Table 4. Comparing these results to the results in the middle panel obtained without price control, we find that the feature coefficient decreases significantly for all products. This pattern aligns with our derivations in Section 2.2, which show that when feature advertising and price are negatively correlated, omitting the price results in an upward bias in the estimated effect of feature advertising. Further, if the estimated price coefficient is negative, even if price is treated as endogenous control, we should expect a smaller bias if the price is included in the regression. The estimates confirm these theoretical results. However, if the price coefficient is underestimated (in absolute terms) due to endogeneity, the regression with the price control but without instruments is still biased upwards – though to a lesser degree than the specification without any price control.

4.2 IV regressions

To address the potential endogeneity of price, we propose an instrument that is based on cross-market correlations in price, similar to commonly used Hausman-type instruments. For each store-week observation, we construct an instrument by computing the chain-specific average price across all stores that belong to the same retailer but are located outside of the focal store’s Designated

Market Area (DMA). The instrument captures chain-wide price fluctuations that are due to changes in input costs or wholesale prices which vary at the chain-level rather than at the level of the individual store. The usual assumption for the instrument to be valid is that demand shocks must be uncorrelated across markets. Because we control for store and week fixed effects, correlated demand shocks can only arise if there are chain-specific temporal demand fluctuations, which are less likely. We note that our instrument is identical the instrument proposed in [DellaVigna and Gentzkow \(2019\)](#), who estimate a two-way fixed effect model, and instrument for price with chain-specific average prices in other markets. They do not include a feature dummy in their regression, because their primary focus is on estimating price elasticities. [Allcott et al. \(2019\)](#) also use a very similar instrument in a cross-sectional setting. They argue that this instrument captures retailer-specific cost advantages in distribution networks.

In the top panel of Table 5 we present the first-stage results. We find that the cost advantage coefficient is close to one and statistically significant. The partial F-statistic of the excluded instrument is large, suggesting that we do not suffer from weak instrument concerns. The coefficient value of close to one suggests a strong correlation in prices across markets, in a manner consistent with [DellaVigna and Gentzkow \(2019\)](#) and [Hitsch et al. \(2021\)](#), who show that many pricing decisions occur at the retail-chain level.

We report results from the second stage of the IV regression in the lower panel of Table 5. We find that the price coefficient becomes, in absolute terms, larger than estimates in the specification without a price instrument. This pattern is consistent with price being set higher in periods of high demand, i.e. $\beta_\xi > 0$ in equation (2). Turning to the featured advertising coefficient, we find that for all three products, the estimated coefficient is significantly smaller than the feature coefficient in the two biased specifications. The magnitude of the bias is substantial. Compared to the regression without price control, the feature coefficient is 3-4 times smaller across the three products. Relative to the endogenous price control regressions, the effect size shrinks by about 40% for two out of three products. These results highlight the importance of properly addressing price endogeneity when estimating the causal effect of feature advertising. We further demonstrate below that similar bias magnitudes are observed across a broader set of product categories.

4.3 Cross-category Results

To extend our analysis, we implement the regressions above across the top 3 products in 7 different product categories, resulting in a total of 21 products. The categories used for this analysis are: Potato Chips, Diapers, Ice Cream, Ketchup, Mayo, Ready-to-Eat Cereal, and Soft Drinks. We implement the same set of regressions that we reported for the mayo category, and plot the estimated coefficients to visualize the patterns that arise across the different categories.

In the top graph in Figure 3, we plot the distribution of the feature coefficient in a regression of price on a feature dummy and fixed effects. These results correspond to the regression results reported for Mayo in the top panel of Table 4. We rank the products by their degree of correlation between feature advertising and price in this and all subsequent graphs. We find a negative impact

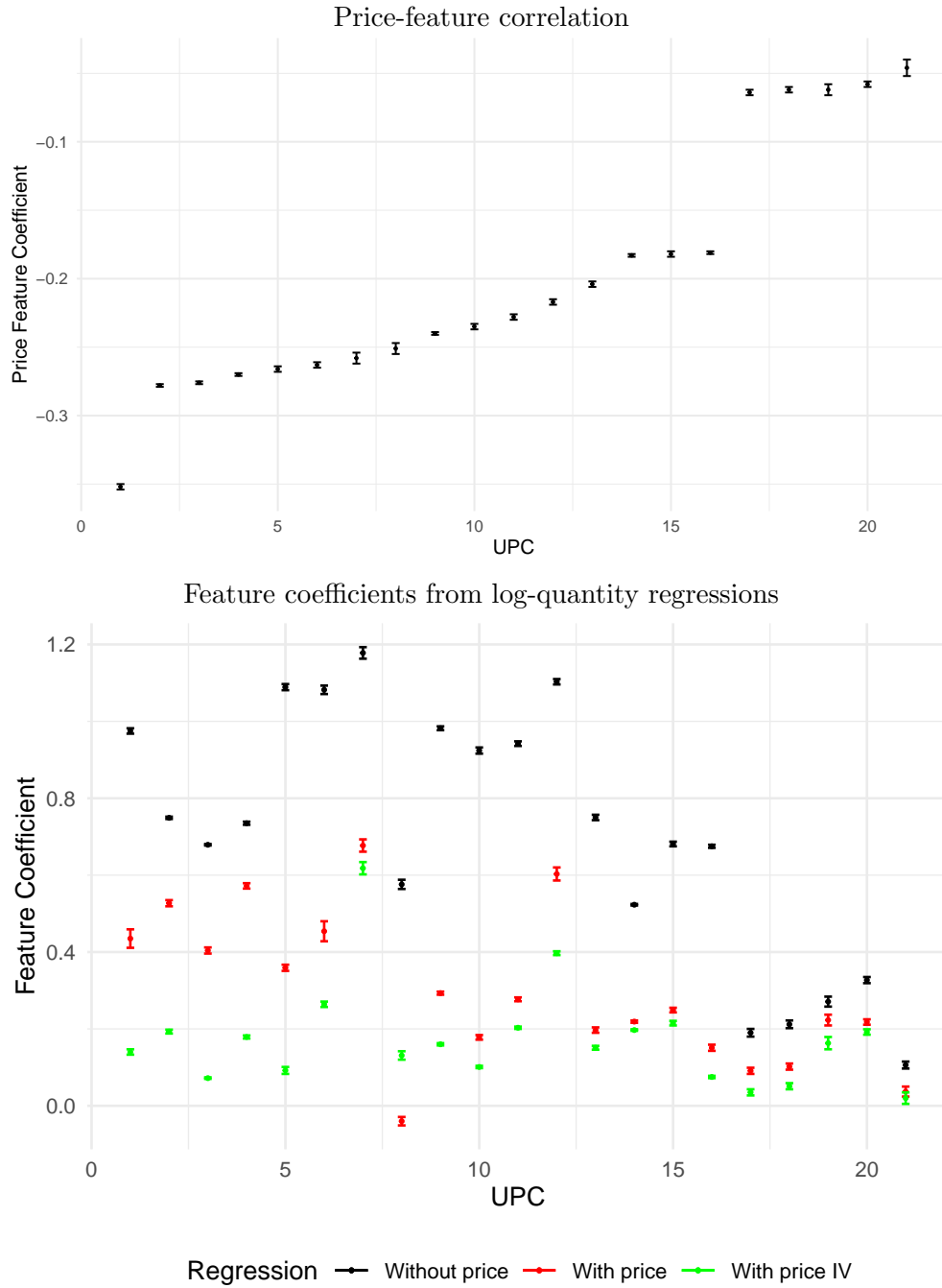


Figure 3: Price / Feature Correlation & Feature Coefficients in Different Specifications. The top panel shows the feature coefficients from regressions of log-price on the feature dummy and fixed effects across products. The bottom panel displays results from regressions of log-quantity sold on the feature dummy and fixed effects, ranked by the size of the price-feature relationship presented in the top panel. We display the feature coefficient when not controlling for price (black), when using price as an endogenous control (red) and when instrumenting price (green).

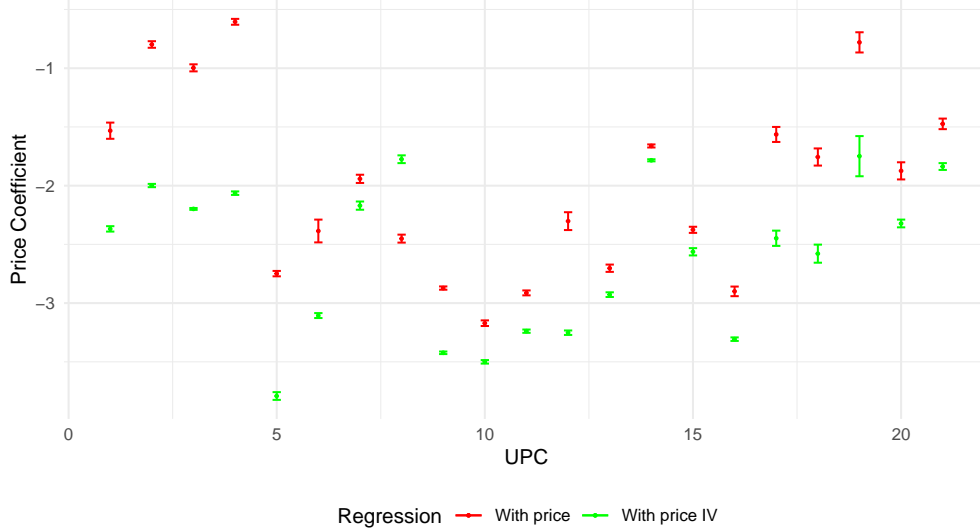


Figure 4: **Price Coefficients with and without Instrument.** Each dot represents the price coefficient from a regression of log-quantity on the feature dummy, log-price and fixed effects, estimated separately by product. Red dots show coefficients from uninstrumented regressions, while green dots reflect those from IV regressions. Products are ordered from left to right by the strength of the price-feature correlation, from most to least correlated.

of feature advertising on (log) price ranging from -0.35 to -0.05. This coefficient corresponds to β_T in the model derivations presented in Section 2. Because feature ads negatively correlate with price, failing to instrument price will lead to a positive bias in the feature coefficient, and hence an overestimation of the impact of a feature. Moreover, larger correlation of the feature with price will lead to a greater bias, both in a regression without price control, and in a regression with endogenous price control. We therefore expect a greater bias for UPCs depicted on left side of the graph.

In the lower graph of Figure 3 we plot the estimated feature coefficients from a set of regressions of quantity on feature advertising. We order products in the same way as in the top graph; that is, by the correlation of feature advertising and price. We report coefficients from a regression without price control (black), with an endogenous price control (red), and when price is instrumented (green). In all comparisons, we treat IV numbers (green) as consistent estimates of true effects. Several patterns in this graph are noteworthy. First, we examine the bias when omitting price controls altogether. As predicted, the bias is larger for products with a stronger correlation between feature advertising and price. For roughly the first half of the UPCs, the bias from omitting price is substantial, and yields an average feature coefficient of around 1.0, whereas regressions using a price instrument lead to an average coefficient of 0.1-0.2. For the final five products, on the right-hand side of the graph, the bias is much smaller because the feature advertising / price correlation is relatively small. In all cases, omitting price leads to upward bias, as the black dots are always above green dots.

Second including price as an endogenous control delivers mixed results. In cases with large correlation of price and feature, the estimated feature coefficient remains significantly biased upwards (red dots are above green dots). However, for several products, especially towards the right half of the graph, the red (endogenous control) and green (instrumented) estimates are similar. This is again consistent with theory as feature and price is less correlated for products toward the right of the graph. For one product, including an endogenous price control leads to a bias in the opposite direction because the price coefficient is overestimated rather than underestimated for the particular product. In all cases, absolute magnitude of the bias is smaller when including price controls.

As we showed earlier, the severity of the bias when using an endogenous price control depends on the bias in the price coefficient (see equation (6)). To illustrate this, we plot the estimated price coefficient from the instrumented and uninstrumented regressions in Figure 4. Consistent with the relative bias reported in Figure 3, the price coefficient is in most cases underestimated (in absolute terms). However, in some instances, such as product 15, the two price coefficients are relatively close together. In this case, price endogeneity is not very severe resulting in a small bias in the estimated effect of feature advertising when controlling for price without instruments.

5 Conclusion

In this paper, we analyze the setting where treatment is (quasi-) randomly assigned, but price and treatment are correlated due to post-treatment price adjustments. We show that the common practice of including price as an (endogenous) control variable leads to a biased treatment effect estimate, where the magnitude of the bias is proportional to the bias in the price coefficient. Although both approaches typically lead to bias, as long as the estimated price coefficient is negative, a regression with an endogenous price controls generates less bias than a regression without any price control. Because bias only occurs when price and treatment are correlated, a simple regression of price on treatment allows the researcher to assess whether post-treatment price adjustments are present. Furthermore, we show that using standard price instruments allows for the unbiased estimation of the treatment effect coefficient. We derive theoretical expressions for the bias in the linear setting. Moreover, we show via simulations that the bias persist across linear and non-linear models. Finally, we illustrate the importance of correctly dealing with endogenous price adjustments using a canonical marketing setting: the estimation of the impact of supermarket feature advertising on demand. We find that the bias from not appropriately controlling for price using instruments can lead to an economically substantial overestimation of the impact of feature advertising on demand.

We hope that this paper raises awareness of the implications of using endogenous price controls and provides researchers with a framework to assess possible biases when using endogenous controls, to test for the presence of price adjustments, and to deal with endogenous prices using instruments.

References

- ACHARYA, A., M. BLACKWELL, AND M. SEN (2016): “Explaining causal findings without bias: Detecting and assessing direct effects,” *American Political Science Review*, 110, 512–529.
- ALLCOTT, H., R. DIAMOND, J.-P. DUBÉ, J. HANDBURY, I. RAHKOVSKY, AND M. SCHNELL (2019): “Food Deserts and the Causes of Nutritional Inequality,” *The Quarterly Journal of Economics*, 134, 1793–1844.
- ANDERSON, E., B. A. MALIN, E. NAKAMURA, D. SIMESTER, AND J. STEINSSON (2017): “Informational rigidities and the stickiness of temporary sales,” *Journal of Monetary Economics*, 90, 64–83.
- ANGRIST, J. D. AND J.-S. PISCHKE (2009): *Mostly harmless econometrics: An empiricist’s companion*, Princeton university press.
- BARAHONA, N., C. OTERO, AND S. OTERO (2023): “Equilibrium effects of food labeling policies,” *Econometrica*, 91, 839–868.
- BERRY, S. (1994): “Estimating discrete-choice models of product differentiation,” *The RAND Journal of Economics*, 242–262.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile Prices in Market Equilibrium,” *Econometrica*, 63, 841–890.
- BLATTBERG, R. C. AND A. LEVIN (1987): “Modelling the effectiveness and profitability of trade promotions,” *Marketing Science*, 6, 124–146.
- BULLOCK, J. G. AND D. P. GREEN (2021): “The Failings of Conventional Mediation Analysis and a Design-Based Alternative,” *Advances in Methods and Practices in Psychological Science*, 4.
- BULLOCK, J. G., D. P. GREEN, AND S. E. HA (2010): “Yes, but what’s the mechanism? (don’t expect an easy answer).” *Journal of Personality and Social Psychology*, 98, 550–558.
- CHINTAGUNTA, P. AND J.-P. DUBÉ (2005): “Estimating an SKU-level Brand Choice Model Combining Household Panel Data and Store Data,” *Journal of Marketing Research*, 42, 368–379.
- CHINTAGUNTA, P., J.-P. DUBÉ, AND K. Y. GOH (2005): “Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models,” *Management Science*, 51, 832–849.
- CINELLI, C., A. FORNEY, AND J. PEARL (2024): “A crash course in good and bad controls,” *Sociological Methods & Research*, 53, 1071–1104.
- DELLAVIGNA, S. AND M. GENTZKOW (2019): “Uniform Pricing in U.S. Retail Chains,” *The Quarterly Journal of Economics*, 134, 2011–2084.
- DUBÉ, J.-P., G. HITSCH, AND P. ROSSI (2018): “Income and Wealth Effects on Private-Label Demand: Evidence From the Great Recession,” *Marketing Science*, 37, 22–53.
- FRISCH, R. AND F. V. WAUGH (1933): “Partial time regressions as compared with individual trends,” *Econometrica: Journal of the Econometric Society*, 387–401.

- GOLI, A., S. MUMMALANENI, P. K. CHINTAGUNTA, AND S. K. DHAR (2022): “Show and sell: studying the effects of branded cigarette product placement in TV shows on cigarette sales,” *Marketing Science*, 41, 1163–1180.
- HARTMANN, W. AND D. KLAPPER (2018): “Superbowl Ads,” *Marketing Science*, 37, 78–96.
- HAYES, A. F. (2017): *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, Guilford publications.
- HENDEL, I. AND A. NEVO (2006): “Measuring the implications of sales and consumer inventory behavior,” *Econometrica*, 74, 1637–1673.
- HITSCH, G. J., A. HORTAÇSU, AND X. LIN (2021): “Prices and promotions in U.S. retail markets,” *Quantitative Marketing and Economics*, 19, 289–368.
- IMAI, K., L. KEELE, D. TINGLEY, AND T. YAMAMOTO (2011): “Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies,” *American Political Science Review*, 105, 765–789.
- JINDAL, P., T. ZHU, P. CHINTAGUNTA, AND S. DHAR (2020): “Marketing-mix response across retail formats: the role of shopping trip types,” *Journal of Marketing*, 84, 114–132.
- JUDD, C. M. AND D. A. KENNY (1981): “Process Analysis: Estimating Mediation in Treatment Evaluations,” *Evaluation Review*, 5, 602–619.
- KONG, X. AND A. RAO (2021): “Do Made in USA Claims Matter?” *Marketing Science*, 40, 731–764.
- KUMAR, N., S. RAJIV, AND A. JEULAND (2001): “Effectiveness of trade promotions: Analyzing the determinants of retail pass through,” *Marketing science*, 20, 382–404.
- MACKINNON, D. P. AND A. G. PIRLOTT (2014): “Statistical Approaches for Enhancing Causal Interpretation of the M to Y Relation in Mediation Analysis,” *Personality and Social Psychology Review*, 19, 30–43.
- MÄNDLI, F. AND M. RÖNKKÖ (2025): “To omit or to include? Integrating the frugal and prolific perspectives on control variable use,” *Organizational Research Methods*, 28, 114–137.
- MONTGOMERY, J. M., B. NYHAN, AND M. TORRES (2018): “How conditioning on posttreatment variables can ruin your experiment and what to do about it,” *American Journal of Political Science*, 62, 760–775.
- PETERSEN, M. L., S. E. SINISI, AND M. J. VAN DER LAAN (2006): “Estimation of Direct Causal Effects,” *Epidemiology*, 17, 276–284.
- PORTER, R. H. (1983): “A Study of Cartel Stability: The Joint Executive Committee, 1880-1886,” *The Bell Journal of Economics*, 14, 301–314.
- ROSENBAUM, P. R. (1984): “The consequences of adjustment for a concomitant variable that has been affected by the treatment,” *Journal of the Royal Statistical Society Series A: Statistics in Society*, 147, 656–666.
- SEILER, S. AND S. YAO (2017): “The impact of advertising along the conversion funnel,” *Quantitative Marketing and Economics*, 15, 241–278.

- SHAPIRO, B., G. HITSCH, AND A. TUCHMAN. (2021): “TV Advertising Effectiveness and Profitability: Generalizable Results from 288 Brands,” *Econometrica*, 89, 1855–1879.
- STREINER, D. L. (2016): “Control or overcontrol for covariates?” *BMJ Ment Health*, 19, 4–5.
- THOMAS, M. (2020): “Spillovers from mass advertising: An identification strategy,” *Marketing Science*, 39, 807–826.

A Bias when Controlling for Price

We first derive the expression for the bias in the price coefficient in a regression of quantity on treatment and price. We let \tilde{p} denote the residual from a regression of price on treatment and δ_T denotes the coefficient when regressing price on treatment. The multivariate regression coefficient on price is given by:

$$\begin{aligned}
 -\hat{\alpha}_p &= \frac{Cov(\tilde{p}, q)}{Var(\tilde{p})} = \frac{Cov(\tilde{p}, -\alpha_1 p + \xi)}{Var(\tilde{p})} \\
 &= -\alpha_p + \frac{Cov(\tilde{p}, \xi)}{Var(\tilde{p})} \\
 &= -\alpha_p + \frac{Cov(p, \xi)}{Var(\tilde{p})} \\
 \alpha_p - \hat{\alpha}_p &= \frac{Cov(p, \xi)}{Var(\tilde{p})}
 \end{aligned}$$

where the second line follows from the fact that $p = \tilde{p} + (\delta_0 + \delta_T Treat)$ and \tilde{p} is uncorrelated with $Treat$. The third line follows because $Treat$ is uncorrelated with ξ due to random assignment of $Treat$.

We can now plug this expression into equation (4):

$$\begin{aligned}
 \alpha_T - \hat{\alpha}_T &= \gamma_p \frac{Cov(p, \xi)}{Var(\widetilde{Treat})} \\
 &= \gamma_p \frac{Cov(p, \xi)}{Var(\tilde{p})} \frac{Var(\tilde{p})}{Var(\widetilde{Treat})} \\
 &= \gamma_p (\alpha_1 - \hat{\alpha}_1) \frac{Var(\tilde{p})}{Var(\widetilde{Treat})} \\
 &= \beta_T \frac{Var(Treat)}{Var(p)} (\alpha_1 - \hat{\alpha}_1) \frac{Var(\tilde{p})}{Var(\widetilde{Treat})} \\
 &= (\alpha_p - \hat{\alpha}_p) \beta_T \frac{Var(Treat)}{Var(\widetilde{Treat})} \frac{Var(\tilde{p})}{Var(p)} \\
 &= (\alpha_p - \hat{\alpha}_p) \beta_T
 \end{aligned}$$

where the fourth line follows because $\gamma_p = Cov(Treat, p)/Var(p)$ and $\beta_T = Cov(Treat, p)/Var(Treat)$ and therefore $\gamma_p = \beta_T Var(Treat)/Var(p)$. The last line follows because the last terms are equal to one.

To see that this term is equal to one, we need to write out the relationship between the variances of treatment and price and their residualized versions. As before, we denote the coefficient when regressing treatment on price by γ_p and the coefficient when regressing price on treatment by δ_T . We can express the variance of the treatment residual as a function of the variance of the treatment

variable as follows:

$$\begin{aligned}
\widetilde{Var(Treat)} &= Var(Treat) - Var(\hat{Treat}) \\
&= Var(Treat) - Var(\gamma_p p) \\
&= Var(Treat) - \gamma_p^2 Var(p) \\
&= Var(Treat) - \frac{Cov(Treat, p)Cov(Treat, p)}{Var(p)} \\
&= Var(Treat) - \frac{Cov(Treat, p)Cov(Treat, p)}{Var(p)Var(Treat)} Var(Treat) \\
&= Var(Treat) - \gamma_p \delta_T Var(Treat) \\
\widetilde{Var(Treat)}/Var(Treat) &= 1 - \gamma_p \delta_T
\end{aligned}$$

and similarly for price:

$$\begin{aligned}
Var(\tilde{p}) &= Var(p) - Var(\hat{p}) \\
&= Var(p) - Var(\delta_T Treat) \\
&= Var(p) - \delta_T^2 Var(Treat) \\
&= Var(p) - \frac{Cov(Treat, p)Cov(Treat, p)}{Var(Treat)} \\
&= Var(p) - \frac{Cov(Treat, p)Cov(Treat, p)}{Var(Treat)Var(p)} Var(p) \\
&= Var(p) - \gamma_p \delta_T Var(p) \\
Var(\tilde{p})/Var(p) &= 1 - \gamma_p \delta_T
\end{aligned}$$

From the two expressions above, it follows that $\frac{Var(\widetilde{Treat})}{Var(Treat)} \frac{Var(\tilde{p})}{Var(p)} = 1$.

The expression in equation (5) can be obtained by plugging the price coefficient bias formula into equation (6):

$$\begin{aligned}
\alpha_T - \hat{\alpha}_T &= (\alpha_p - \hat{\alpha}_p) \beta_T \\
&= \beta_T \frac{Cov(p, \xi)}{Var(\tilde{p})} \\
&= \beta_T \beta_\xi \frac{Var(\xi)}{Var(\tilde{p})}
\end{aligned}$$