The Paper of How: Estimating Treatment Effects Using the Front-Door Criterion

Marc F. Bellemare* Jeffrey R. Bloem†

August 16, 2019

Abstract

Empirical social science nowadays consists largely of attempts at answering questions of the form "Does X cause Y?" As a result, social scientists rely on a number of empirical techniques aimed at disentangling causal relationships from mere correlations. One such technique is Judea Pearl's (1995, 2000) front-door criterion, which relies for identification on the presence of a single, strictly exogenous mechanism on the causal path between the treatment and outcome. Social scientists in general—and economists in particular—have been resistant to the idea of adding the front-door criterion to the standard empirical toolkit, largely due to the difficulty posed by finding the required mechanism. To help overcome that resistance, we first explain how to use the front-door criterion in a regression context. We then present three empirical illustrations of the front-door criterion. Finally, and most importantly, we look at what happens when some of the assumptions underpinning the front-door criterion are violated.

Keywords: Front-Door Criterion, Causal Inference, Causal Identification, Treatment Effects

JEL Codes: C13, C18

*Corresponding Author. Northrop Professor, Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, Saint Paul, MN 55108, Email: mbellema@umn.edu.

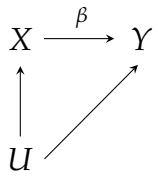
†Ph.D. Candidate, Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, Saint Paul, MN 55108, Email: bloem023@umn.edu.

1 Introduction

Empirical social science research nowadays consists largely of attempts at answering questions of the form "Does X cause Y?" Consequently, social scientists rely on a number of empirical techniques aimed at disentangling causal relationships from mere correlations. The fact that one cannot usually infer a causal relationship from a mere correlation is referred to as the "identification problem."

For a dichotomous treatment variable X and a continuous outcome Y, Figure I graphically represents the identification problem, where the average treatment effect (ATE), $\beta = E(Y|X = 1) - E(Y|X = 0)$, is not identified because of the presence of unobserved confounders U.

FIGURE I: The Identification Problem

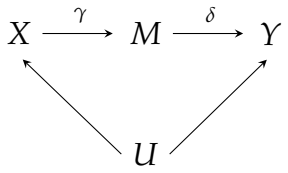


For instance, suppose we observe a correlation between someone consuming red wine X and experiencing migraines Y. It is possible that drinking red wine causes migraines. It is also possible, however, that a third, unobserved factor U—stress, for instance—causes one to both drink red wine and experience migraines. In order to test whether drinking red wine causes migraines, applied researchers have to use one of the several techniques available to social scientists for identifying causal relationships. In the ideal case, the researcher conducts an experiment by randomly assigning each observation i to the control (i.e., $X_i = 0$, drinks no red wine) or treatment (i.e., $X_i = 1$, drinks red wine) groups, which frees X from the influence of U and allows for credible causal identification.

Not all questions, however, can be answered with an experiment. Practical limitations from implementation challenges to ethical concerns often prevent the use of experiments to answer important research questions, and so in the absence of experimental data, most

empirical social scientists rely on quasi-experimental techniques. All of those techniques (e.g., instrumental variables, difference-in-differences, regression discontinuity) exploit a specific feature of the data at hand to get as close as possible to a situation where X is experimentally assigned, and thus not affected by the unobserved confounder U.

FIGURE II: The Front-Door Criterion



One such quasi-experimental technique is Pearl's (1995, 2000) front-door criterion (FDC), an empirical method which was brought to the general public's attention in Pearl and Mackenzie's (2018) *The Book of Why*, and which is represented in Figure II. The idea behind the FDC is simple. Suppose one is interested in identifying the causal relationship flowing from X to Y in Figure I. The unobserved confounder U obviously prevents identifying that causal relationship, but if one can find a mechanism M which (i) lies on the causal path between X and Y, (ii) is the only such mechanism, and (iii) is not affected by the unobserved confounder U, then one can identify the causal relationship flowing from X to Y.

The intuition behind this is as follows. Breaking down Figure II into two parts yields Figure III, which represents to causally identified relationships. On the left, in panel (a), the causal relationship γ from X to M is identified because the unobserved confounder U affects X but does not affect M. On the right, in panel (b), the causal relationship δ is identified because the unobserved confounder affects Y but does not affect M. Combining the two parts of the FDC—the ATE γ of X on M, and the ATE δ of M on Y—by multiplying them by one another allows recovering β, the ATE of X on Y.

For instance, it may be that we are not able to randomly assign whether someone consumes red wine on a given day (i.e., X), but we can randomly how much red wine this

FIGURE III: The Front-Door Criterion as the Product of Two Identified Causal Relationships

person drinks (i.e., M) when they decide to consume red wine. In that case, assuming perfect compliance, the FDC can credibly estimate the causal effect of red wine consumption on migraines (i.e., Y).

Social scientists in general—and economists in particular—have been resistant to incorporating the FDC in their standard empirical toolkit. This resistance largely stems from the fact that, for any application wherein a researcher is interested in identifying the causal impact of X on Y, finding a mechanism M which satisfies the criteria laid out above can be very difficult.[1] In his writings on the FDC, for instance, Judea Pearl always gives the same canonical example where X is a dummy variable for whether one smokes, Y is a dummy variable for whether one develops lung cancer, and M is the accumulation of tar in one's lungs, which may be due to either to one smoking or to the conditions of one's environment (Pearl 1995; Pearl 2000; Pearl and Mackenzie 2018). Consequently, the only extant published social science applications of the FDC are by Glynn and Kashin (2017; 2018), who respectively show that the front-door adjustment can be useful even when the assumptions required for the FDC do not hold, and develop a front-door difference-in-differences estimator as an extension of the FDC.[2]

The contribution of this paper is thus threefold. First, because linear regression remains the workhorse of empirical social science, and because an explanation of how to use the FDC in a regression context has so far been lacking in the literature, we explain

[1] A well-articulated example of economists' skepticism regarding the usefulness of the front-door criterion can be found in section 2.7 of Imbens (2019).

[2] In Glynn and Kashin (2018) the authors apply the FDC approach to estimating the effects of attending a job training program on earnings. In Glynn and Kashin (2017) the authors also apply the FDC Difference-in-Differences approach to evaluate the effect of an early in-person voting program on voting turnout.

how to use the FDC in the context of linear regression. Second, we present three examples of the FDC in practice: one using simulated data, and two using real-world data where the randomly varying treatment intensity serves as the mechanism M—one from a randomized controlled trial, and one from a lab experiment. Lastly, and perhaps most importantly for applied econometricians, we explore departures from Pearl's ideal FDC scenario. Specifically, we look at what happens when (i) there are multiple mechanisms, some of which may be omitted from the FDC estimation, (ii) the assumption of strict exogeneity of M is violated, and (iii) the treatment is completely defined by the mechanism.

The remainder of this paper is organized as follows. In section 2, we present the theory behind the FDC and a "how-to" for those applied economists who may wish to broaden their empirical toolkit by adding to it the FDC. Section 3 presents thee empirical illustrations: one based on simulated data, and two based on the real-world data in Beaman et al. (2013) and Bellemare et al. (2019). In section 4, we explore departures from some of the assumptions underpinning the FDC. We conclude in section 5 by offering practical recommendations for using the FDC in empirical research.

2 The Front-Door Criterion: Theory and Practice

We begin this section with a brief presentation of the FDC estimand as derived in Pearl (1995, 2000). We then offer our first contribution, which is that of explaining how to use the FDC in practice in a regression context.

2.1 Theory

We are interested in estimating $E[Y|X]$, the ATE of X on Y in Figure I above. Recall that with observational data, this task is complicated by the presence of unobserved confounders, U, which give rise to the identification problem. Given the validity of a number of identifying assumptions, however, the FDC approach pictured in Figure II allows for

unbiased estimates of causal effects.

As discussed in Pearl (1995, 2000), the FDC requires that there exists a variable, M, that satisfies the following conditions relative to X and Y:

> (i) *M intercepts all directed paths from X to Y.* This implies that the only way in which X influences Y is through M. In Figure II, this means that there should be no arrows bypassing M between X and Y.
>
> (ii) *There is no back-door path between X and M.* This implies that the relationship between X and M is not confounded by unobserved confounders, i.e., that the coefficient $\gamma = E[M|X]$ in Figure II is identified.
>
> (iii) *Every back-door path between M and Y is blocked by X.* This implies that, conditional on X, the relationship between M and Y is not confounded by unobserved confounders, i.e., the coefficient $\delta = E[Y|M, X]$ in Figure II is identified.

As in Pearl (1995), we derive the FDC estimand in three steps, with the goal being to compute $P(Y|do(X))$ with observable variables, where $do(X)$ denotes X when that variable is randomly assigned. Therefore, $P(Y|do(X))$ represents the causal effect of X on Y. This should be contrasted with $P(Y|X)$, which may not represent the causal effect of X on Y due to the presence of the unobserved confounder U. Because observing $do(X)$ is unlikely outside of a randomized experiment, our goal here is to restate $P(Y|do(X))$ using only observational variables.

The first step is to compute $P(M|do(X))$. Here we make use of condition (ii), i.e., there is no back-door path between X and M, and so the relationship between X and M is identified. When that condition holds, we can write

$$P(M|do(X)) = P(M|X), \tag{1}$$

given that in this case, the unobserved confounder U affecting X but not M makes the two sides of Equation 1 equivalent.

The second step is to compute $P(Y|do(M))$. Here we cannot set $do(M) = M$ because there is a back-door path from M to Y, via X. To block this path we can make use of condition (iii), i.e., every back-door path between between M and Y is blocked by X. In that case, by controlling for and summing over all values X_i of X, we can write

$$P(Y|do(M)) = \sum_X P(Y|X, do(M)) \times P(X|do(M)) \tag{2}$$

where the right-hand-side of Equation 2 involves two expressions involving $do(M)$. The second term on the right-hand-side of Equation 2 can be reduced to $P(X)$ because, as stated by condition (i), M intercepts all directed paths from X to Y. Put differently, if M were randomly assigned then it would not influence the probability of X. The first term on the right-hand-side of Equation 2 can be expressed as $P(Y|X, M)$ because, as stated by condition (iii), X blocks all back-door paths from M to Y. This implies that conditional on X, the relationship between M and Y is identified. Therefore, we can write

$$P(Y|do(M)) = \sum_X P(Y|X, M) \times P(X). \tag{3}$$

The third and last step is to combine the two effect estimates, $P(M|do(X))$ from Equation 1 and $P(Y|do(M))$ from Equation 2, in order to compute $P(Y|do(X))$—the causal effect of X on Y.

To start, we express $P(Y|do(X))$ in terms of $do(X)$ by controlling for and summing over all values of M_i of M. This allows us to write

$$P(Y|do(X)) = \sum_M P(Y|M, do(X)) \times P(M|do(X)). \tag{4}$$

Condition (iii) allows us to rewrite M as $do(M)$ in the first term on the right-hand-side of Equation 4. Because every back-door path between Y and M is blocked by X, then the variation in M is exogenous. Additionally, as stated by condition (i), M intercepts all paths from X to Y and so we can remove $do(X)$ from the first term on the right-hand side

of Equation 4. Put differently, the causal effect between M and Y is independent of X, so we can rewrite the first term on the right-hand side of Equation 4 as

$$P(Y|M, do(X)) = P(Y|do(M), do(X)) = P(Y|do(M)). \tag{5}$$

Recall that Equation 3 states that $P(Y|do(M)) = \sum_X P(Y|X, M) \times P(X)$ and Equation 1 states that $P(M|do(X)) = P(M|X)$. Therefore, plugging Equation 3 into Equations 4 and 5, and plugging Equation 1 into Equation 4 gives us the FDC estimand as originally derived by Pearl (1995). That estimand is such that

$$P(Y|do(X)) = \sum_M P(M|X) \times \sum_{X'} P(Y|X', M) \times P(X'). \tag{6}$$

Conceptually, and as discussed in Hernán and Robins (2019), the FDC approach works by first estimating the effect of X on M, and then estimating the effect of M on Y holding X constant. Both of these effects are unbiased because nothing confounds the effect of X on M and X blocks the only back-door path between M on Y. Multiplying these effects by one another yields the FDC estimand.

Pearl (2000) makes one more assumption. This assumption states that $P(X_i|M_i) > 0$. This implies that no matter the value of the mechanism is for unit i, that unit has to have a nonzero probability of getting treated. Among other things, this means that the treatment X_i cannot be entirely defined by the mechanism M_i. In the canonical example of the relationship between smoking X and lung cancer Y, this assumption implies that the amount of tar in the lungs of smokers M must be the the result not only of smoking, but also of other factors (e.g., exposure to environmental pollutants), and that tar be absent from the lungs of some smokers (say, because of an extremely efficient tar-rejecting mechanism). We will discuss this assumption in more detail in Section 4, but the good news is that it can be can be directly validated with data.

2.2 Practice

In this sub-section we discuss how to empirically estimate treatment effects using the FDC approach. When the necessary conditions for the FDC to identify a treatment effect hold, one way to estimate treatment effects is using the following approach. Let

$$M_i = \kappa + \gamma X_i + \omega_i \tag{7}$$

and

$$Y_i = \lambda + \delta M_i + \phi X_i + \nu_i. \tag{8}$$

In Equation 7, following condition (ii) which states that there is no back-door path between X and M, the relationship between X and M is identified, since $Cov(X, \omega) = 0$. In Equation 8, Y_i is the outcome variable, which is related to X_i only through M_i. In this case, following conditions (i) and (iii) which together imply that M intercepts all paths between X and Y, and X blocks all back-door paths between M and Y, the relationship between M and Y conditional on X is identified, since $Cov(M, \nu) = 0$. Therefore, estimating Equations 7 and 8 and multiplying the coefficient estimates $\hat{\delta}$ and $\hat{\gamma}$ by each other estimates β.

More formally, one can write

$$ATE_{FDC} = E[Y|do(X)] = E[\hat{\delta} \times \hat{\gamma}] = \beta. \tag{9}$$

The conditions under which equation 9 must hold are derived as follows. The definition of β, from Figure I, is:

$$\beta = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \tag{10}$$

By subtracting the mean of Y_i in equation 8 and substituting this into equation 10 yields the following augmented definition of β.

$$\beta = \frac{\sum(X_i - \bar{X})(\delta_1(M_i - \bar{M}) + \phi(X_i - \bar{X}) + (\nu_i - \bar{\nu}))}{\sum(X_i - \bar{X})^2} \tag{11}$$

Similarly, by subtracting the mean of M_i in equation 7 and substituting this into equation 11 yields the following augmented definition of β, which is now in terms of X_i and the error terms from equations 7 and 8.

$$\beta = \frac{\sum(X_i - \bar{X})(\delta(\gamma(X_i - \bar{X}) + (\omega_i - \bar{\omega})) + \phi(X_i - \bar{X}) + (\nu_i - \bar{\nu}))}{\sum(X_i - \bar{X})^2} \tag{12}$$

Equation 12 reduces to the following expression

$$\beta = \delta \times \gamma + \delta \frac{Cov(X, \omega)}{Var(X)} + \phi + \frac{Cov(X, \nu)}{Var(X)} \tag{13}$$

If $Cov(X, \omega) = 0$, $Cov(X, \nu) = 0$, and $\phi = 0$, then $\beta = \delta \times \gamma$. The exogeneity of M leads to $Cov(X, \omega) = 0$ and $Cov(X, \nu) = 0$. This assumption is implied by conditions (ii) and (iii) which together state that there is no back-door path between X and M, and that every back-door path between M and Y is blocked by X. Finally, $\phi = 0$ indicates that there is no direct effect of X on Y conditional on both M and the unobserved confounder U. This "no direct effect" assumption is implied by condition (i)—that M intercepts all directed paths from X to Y.

In section 4, we will examine the consequences of the failure of some of these assumptions. Specifically, we will look at what happens when $Cov(X, \omega) \neq 0$ and $Cov(X, \nu) \neq 0$ (i.e., when M is not strictly exogenous) and when $\phi \neq 0$ (i.e., when there is a direct effect of X and Y). Both of these cases will lead to biased estimates of the ATE defined by the formula given in Equation 13.

3 Empirical Illustration

We first show empirical results using simulated data. We then show results using data from a randomized controlled trial conducted by Beaman et al. (2013) in Mali. Finally, we show results using data from a lab experiment conducted by Bellemare et al. (2019) at two experimental labs in the US in the field in Peru.

3.1 Simulation Results

Our simulation setup is as follows. Let $U_i \sim N(0,1)$, $Z_i \sim U(0,1)$, $\epsilon_{Xi} \sim N(0,1)$, $\epsilon_{Mi} \sim N(0,1)$, and $\epsilon_{Yi} \sim N(0,1)$ for a sample size of $N = 100,000$ observations. Then, let

$$X_i = 0.5U_i + \epsilon_{Xi}, \tag{14}$$

$$M_i = Z_i X_i + \epsilon_{Mi}, \tag{15}$$

and

$$Y_i = 0.5M_i + 0.5U_i + \epsilon_{Yi}. \tag{16}$$

This fully satisfies Pearl's (1995, 2000) three criteria for the FDC to be able to estimate the average treatment effect of X on Y, viz. (i) the only way in which X influences Y is through M, (ii) the relationship between M and X is not confounded by U, since U only affects X and not M, and (iii) conditional on X, the relationship between M and Y is not confounded by U. By substituting Equation 14 into Equation 15, and then Equation 15 into Equation 16, it should be immediately obvious to the reader that the true ATE is equal to 0.250 in our simulations.

To show that the FDC estimates the ATE of X on Y, we estimate two specifications. The first specification, which we refer to as our benchmark specification because it generates an unbiased estimate of the ATE by virtue of controlling for the unobserved confounder U, estimates

$$Y_i = \alpha_0 + \beta_0 X_i + \zeta_0 U_i + \epsilon_{0i}, \tag{17}$$

where, because both X_i and U_i are included on the right-hand-side, $E(\hat{\beta}_0) = \beta$, i.e., the true ATE.

The second specification, which we refer to as our front-door specification, estimates

$$M_i = \kappa_0 + \gamma_0 X_i + \omega_{0i} \tag{18}$$

$$Y_i = \lambda_0 + \delta_0 M_i + \phi_0 X_i + \nu_{0i} \tag{19}$$

where the unobserved confounder U_i does not appear anywhere, but because the necessary assumptions for the FDC to identify the ATE hold, $E(\hat{\gamma}_0 \cdot \hat{\delta}_0) = \beta$, i.e., the true ATE.

Lastly, we estimate a naïve specification, one that is similar to the benchmark specification in Equation 17, but which fails to control for the presence of the unobserved confounder.

Column 1 of Table I shows estimation results for Equation 17, our benchmark specification. Column 2 shows estimation results for our naïve specification. Columns 3 and 4 show estimation results respectively for the front-door specification in Equations 18 and 19, respectively. The line labeled "Estimated ATE" shows estimates of the ATE for each of those three specifications. Unsurprisingly, the estimates of the ATE in columns 1 and 2 differ markedly, as the former controls for U_i but the latter does not: $\hat{\beta}$ is equal to 0.252 in the benchmark case, but it is near double that at 0.454 in the naïve case.

Given the derivations above, it should also come as no surprise that the estimate of the ATE generated by multiplying the coefficient on treatment in column 3 by the coefficient on the mechanism in column 4 is equal to 0.254. Assuming the ATE in column 1 is not correlated with the ATE computed from columns 3 and 4, the benchmark and front-door ATEs are statistically identical. In both cases, the estimated ATE is not statistically

TABLE I: Simulation Results

	Benchmark	Naïve	Front-Door		Direct Effect
Variables	Y	Y	M	Y	Y
	(1)	(2)	(3)	(4)	(5)
Treatment (X)	0.252***	0.454***	0.507***	0.200***	-0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
Mechanism (M)	–	–	–	0.502***	0.500***
				(0.003)	(0.003)
Confounder (U)	0.499***	–	–	–	0.501***
	(0.004)				(0.004)
Intercept	-0.004	-0.005	-0.004	-0.003	-0.003
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
Estimated ATE	0.252***	0.454***	0.254***		–
	(0.004)	(0.003)	(0.002)		
Observations	100,000	100,000	100,000		100,000

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The front-door equations in columns (3) and (4) are estimated by seemingly unrelated regressions. The standard error for the front-door ATE is estimated by the delta method.

different from its true value of 0.250.

Finally, column 5 in Table 1 serves an illustrative purpose. It shows that conditional on the mechanism (M) and the unobserved confounder (U), the coefficient on the treatment (X) is statistically indistinguishable from zero. This result highlights the "no direct effect" assumption that is implied by the FDC conditions in Figure I and by equation 13.

3.2 Real-World Results

This section illustrates the FDC using the results of two experimental studies, one by Beaman et al. (2013), and one by Bellemare et al. (2019).[3] In Table II, we replicate results from Beaman et al. (2013), who conduct a randomized controlled trial with rice farmers in Mali. Starting from the full sample, units of observations are either assigned to a treat-

[3]In both of the examples presented in this section, treatment is randomly assigned, and so there is no unobserved confounder U affecting both X and Y. In addition, Pearl's assumption that $P(X_i|M_i) > 0$ is violated in both the examples because in cases where the mechanisms are equal to zero, the likelihood that an observation will receive treatment is equal to zero. As a consequence of those two facts, the estimation procedure discussed in section 2 has to be modified slightly. We explain in section 4.3 how to do so.

TABLE II: Empirical Illustration — Rice Production and Fertilizer Use in Mali

	Use of Fertilizer		Fertilizer Quantity		Fertilizer Expenses	
	(1)	(2)	(3)	(4)	(5)	(6)
Benchmark	0.639***		27.24***		-2,717.1***	
	(0.033)		(3.568)		(464.6)	
Front-Door		0.603***		26.64***		-2,605.3***
		(0.030)		(3.002)		(389.7)
Observations	378	378	378	378	373	373

Notes: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All columns include the same control variables as in Beaman et al. (2013). Columns (1), (3), and (5) represent benchmark OLS estimates of the ATEs of receiving fertilizer on the outcome variables. The estimates in columns (1), (3), and (5) replicate the findings of Beaman et al. (2013) except that the original study differentiates between two treatment groups defined by intensity of treatment. Columns (2), (4), and (6) represent seemingly unrelated regression estimates of the front-door criterion ATEs. Standard errors in columns (2), (4), and (6) are estimated by the delta method.

ment group or a control group, with treated units receiving fertilizer and control units receiving no fertilizer.

Here, we exploit as a mechanism the fact that treatment intensity varies at random within the treatment group to illustrate the FDC in practice. Treatment intensity varies at random, with about half of the treatment-group observations receiving half of the prescribed amount of fertilizer, the remainder of the treatment-group observations receiving the full prescribed amount of fertilizer, and the control-group observations receiving none of the prescribed amount of fertilizer. As one would expect from the derivations in Section 2, the results in Table II show that the ATEs obtained by the FDC are all statistically indistinguishable from the benchmark ATEs. The results show that receipt of free fertilizer leads to increases in the use of fertilizer at both the extensive and intensive margins and reduces fertilizer expenses.

Table III also illustrates the FDC using real-world data using data from Bellemare et al.'s (2019) lab experiment aimed at testing Sandmo's (1971) theoretical prediction that output price risk makes risk-averse producers hedge against price risk by producing less than they do in situations of price certainty. Here, subjects are put in the role of producers facing a price that is either certain or uncertain, with the amount of price uncertainty—

the standard deviation of the price distribution—varying at random in the subset of cases where the price is risky. Specifically, Y denotes output quantity chosen by a given subject in a given round ex ante of the realization of the price, X denotes whether the price is certain ($X = 0$) or uncertain ($X = 1$), and M measures treatment intensity, viz. how uncertain the output price, or the standard deviation of the price distribution, which is chosen at random from four distinct values.

The results in Table III show that the ATEs obtained by the FDC are statistically indistinguishable from the benchmark ATEs. These results suggest that in the whole sample, subjects increase their production by a quarter of one unit in response to price risk. This masks some heterogeneity, however: In US labs, subjects do not change their production in response to price risk, but in Peru, farmers increase their production by 0.6 units in response to price risk.

A few remarks are in order. First, the real-world results in this section are most useful for highlighting the potential of the FDC approach in estimating treatment effects in settings where each of the conditions hold. Of course, since both Beaman et al. (2013) and Bellemare et al. (2019) have the benefit of experimentally assigning treatment, the FDC approach is not necessary to estimate treatment effects in either context.

Second, in these real-world experimental cases, there is no need to condition on the treatment variable (i.e., X) when estimating the effect of the mechanism (i.e., M) on the outcome (i.e., Y) since the random assignment of treatment already removes any back-door path between Y and M. In fact, needlessly conditioning on the treatment variable in an experimental setting leads to bias in the front-door estimate, due to violating the assumption that $P(X_i|M_i) > 0$. We return to this phenomenon in section 4.3.

TABLE III: Empirical Illustration — Price Risk and Production in the Lab

	Whole Sample		US (Lab, Students)		Peru (Field, Farmers)	
	(1)	(2)	(3)	(4)	(5)	(6)
Benchmark	0.252**		0.082		0.591***	
	(0.123)		(0.135)		(0.227)	
Front-Door		0.229**		0.04		0.603***
		(0.113)		(0.124)		(0.209)
Observations	2,339	2,339	1,419	1,419	920	920

Notes: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. No control variables are included. Columns (1), (3), and (5) represent benchmark OLS estimates of the ATEs on production of facing a risky price rather than a certain price. Columns (2), (4), and (6) represent seemingly unrelated regression estimates of the front-door criterion ATEs. Standard errors in columns (2), (4), and (6) are estimated by the delta method.

4 Departures from the Ideal Case

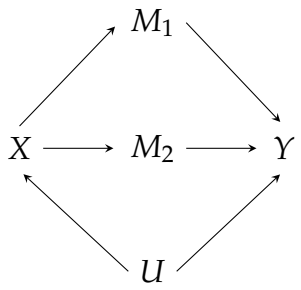
Having discussed how one can use the FDC to estimate ATEs both in theory and in practice in section 2, and having illustrated the use of the FDC to estimate ATEs using both simulation and real-world data in section 3, we now turn to investigate what happens when some of the assumptions required for the FDC to identify an ATE fail to hold. To do so, we look in turn at what happens with multiple mechanisms, when the mechanism is no longer strictly exogenous, and when the treatment is totally defined by the mechanism.

4.1 Multiple Mechanisms

Pearl's (1995, 2000) canonical treatment of the front-door criterion assumes that M is a single variable, and not a vector of mechanism variables. Consequently, in the empirical examples in section 3, we considered cases where the mechanism, M, is defined by a single variable rather than a vector. In this sub-section we consider how to implement a case where we have multiple mechanisms.

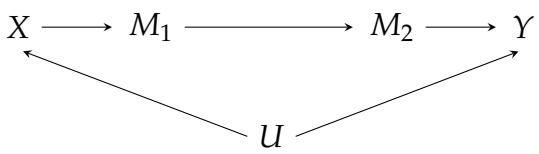
There are two basic cases in which we can imagine multiple mechanisms. Of course, one can imagine more complicated cases that combine these two basic cases. For the illustrative purposes, however, we will examine these two cases separately. In the first case, as shown in Figure IV, the multiple mechanisms are independent from each other.

FIGURE IV: Multiple Mechanisms—Case 1



Specifically, a path flows from X to both M_1 and M_2, and additionally, a path flows from both M_1 and M_2 to Y. In this case, M_1 and M_2 together intercept all directed paths from X to Y and meet the requirement of condition (i). By simply examining Figure IV it is clear that omitting either M_1 or M_2 from the estimation will violate condition (i), since the single mechanism does not intercept all directed paths from X to Y.

FIGURE V: Multiple Mechanisms—Case 2



In the second case, as shown in Figure V, the multiple mechanisms both lie on the same path between X and Y. Specifically, a path flows from X to M_1, from M_1 to M_2, and finally from M_2 to Y. In this case, either M_1 or M_2 intercept all directed paths from X to Y and meet the requirement of condition (i). In contrast to the previous case, omitting either M_1 or M_2 from the estimation will not violate condition (i), since both mechanisms individually intercept all directed paths from X to Y. Therefore the FDC approach will recover the ATE when using either only M_1, only M_2, or both M_1 and M_2 as mechanisms in the FDC estimation. This point should be obvious based on conceptual reasoning, but a simulation showing this result can be found in the appendix.

We now show simulation results that demonstrate the consequences of multiple mechanisms of the sort illustrated in Figure IV, where multiple mechanisms lie on different

paths from X to Y.

Our simulation setup is as follows. Let $U_i \sim N(0,1)$, $\epsilon_{Xi} \sim N(0,1)$, $Z_{1i} \sim U(0,1)$, $Z_{2i} \sim U(0,1)$, $\epsilon_{M1i} \sim N(0,1)$, $\epsilon_{M2i} \sim N(0,1)$, and $\epsilon_{Yi} \sim N(0,1)$ for a sample size of $N = 100,000$ observations. Then, let

$$X_i = 0.5U_i + \epsilon_{Xi}, \tag{20}$$

$$M_{1i} = Z_{1i}X_i + \epsilon_{M1i}, \tag{21}$$

$$M_{2i} = Z_{2i}X_i + \epsilon_{M2i}, \tag{22}$$

and

$$Y_i = 0.5M_{1i} + 0.5M_{2i} + 0.5U_i + \epsilon_{Yi}. \tag{23}$$

As illustrated in Figure IV, this fully satisfies Pearl's (1995, 2000) three criteria for the FDC to be able to estimate the average treatment effect of X on Y. By substituting Equation 20 into Equations 21 and 22, and then substituting these equations into Equation 23, it should be immediately obvious to the reader that the true ATE is equal to 0.500 in our simulations.

Similar to the previous simulation analysis, we estimate two specifications. The first specification estimates

$$Y = \alpha_1 + \beta_1 X_i + \zeta_1 U_i + \epsilon_{1i}, \tag{24}$$

where, because both X and U are included on the right-hand-side, $E(\hat{\beta}_1) = \beta$, i.e., the true ATE.

The second specification estimates

$$M_{1i} = \kappa_1 + \gamma_1 X_i + \omega_{1i} \tag{25}$$

$$M_{2i} = \pi_1 + \rho_1 X_i + \xi_{1i} \tag{26}$$

$$Y_i = \lambda_1 + \delta_1 M_{1i} + \tau_1 M_{2i} + \phi_1 X_i + \nu_{1i} \tag{27}$$

where the unobserved confounder U does not appear anywhere. The small difference in the case of multiple independent mechanisms is the true ATE is calculated by adding two products together, $E[(\hat{\gamma}_1 \cdot \hat{\delta}_1) + (\hat{\rho}_1 \cdot \hat{\tau}_1)] = \beta$.

Column 1 of Table IV shows our benchmark estimation results for Equation 24. Column 2 shows estimation results for the naïve version of Equation 24 which omits the unobserved confounder U. Columns 3, 4, and 5 show FDC estimation results using the specification outlined in Equations 25 to 27, respectively. Again, the estimates of the ATE in columns 1 and 2 are quite different, and the ATE estimate is equal to 0.501 in the benchmark case, but it is much larger, at 0.703, in the naïve case.

Given the derivations above, it should come as no surprise that the estimate of the ATE generated by the FDC approach accurately estimates the ATE. The FDC approach first multiplies the coefficient on X in column 3 by the coefficient on M_1 in column 5. Next, the FDC approach multiplies the coefficient on X in column 4 by the coefficient on M_2 in column 5. Finally, these two products are summed to estimate the ATE. Assuming the ATE in column 1 is not correlated with the ATE computed from columns 3 through 5, the two ATEs are statistically identical. In both cases, the estimated ATE is not statistically different from its true value of 0.500. Finally, in column 6, the direct effect of treatment conditional on M_1, M_2, and U is statistically indistinguishable from zero.

More interesting, however, is investigating and interpreting estimates when we erroneously omit one of the mechanisms (say, for example, M_2) from the FDC estimation. In this case, we no longer can correctly assume no "direct effect" of X on Y since there is a directed path independent of M_1 via M_2. This violates condition (i) of Pearl (1995, 2000).

Given the formula for bias in the FDC estimation approach, shown in equation 13, in the presence of a non-zero "direct effect" when omitting a mechanism, the estimate of

the ATE will be biased by $-\phi$. Columns 7 through 9 in Table IV domonstrate this fact. When we omit M_2 from the FDC estimation the estimated ATE (shown in columns 7 and 8) is 0.246, quite a bit smaller than the true ATE. Column 9 shows that the "direct effect" is 0.254. Indeed, as implied by equation 13, if we add the biased FDC estimate from columns 7 and 8 to the coefficient on treatment from column 9 we recover the the true ATE.

TABLE IV: Simulation Results—Multiple Mechanisms, Case 1

	Benchmark	Naïve	Front-Door			Direct Effect	Biased Front-Door		Direct Effect
Variables	Y	Y	M_1	M_2	Y	Y	M_1	Y	Y
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment (X)	0.501***	0.703***	0.497***	0.502***	0.204***	0.001	0.497***	0.457***	0.254***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
Mechanism (M_1)	–	–	–	–	0.498***	0.500***	–	0.495***	0.496***
					(0.003)	(0.003)		(0.004)	(0.003)
Mechanism (M_2)	–	–	–	–	0.499***	0.499***	–	–	–
					(0.003)	(0.003)			
Confounder (U)	0.498***	–	–	–	–	0.501***	–	–	0.500***
	(0.004)					(0.004)			(0.004)
Intercept	-0.002	-0.003	-0.005	0.002	-0.002	-0.004	-0.005	-0.001	0.001
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Estimated ATE	0.501***	0.703***		0.498***		–	0.246***		–
	(0.004)	(0.004)		(0.003)			(0.002)		
Observations	100,000	100,000		100,000		100,000	100,000		100,000

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The front-door equations in columns (3) and (4) are estimated by seemingly unrelated regressions. The standard error for the front-door ATE is estimated by the delta method.

The foregoing shows the consequences of omitting a mechanism when using the FDC approach to estimate the ATE. With that said, the effect estimated in the biased FDC estimation in columns 7 and 8 of Table IV can be interpreted as the "indirect effect" of X on Y via M_1 and independent of M_2 (Imai et al. 2010, Acharya et al. 2016). In the literature on causal mediation analysis, the total causal effect is framed as the aggregation of both the direct and indirect effects (Imai et al. 2011). The ability to relax condition (i)—that the M intercepts all paths from X to Y—may allow for more credible and useful applications of the FDC approach in assisting applied researchers estimate mediation effects. Of course, whether or not the indirect effect is a parameter of interest for applied researchers will ultimately depend on the specific empirical application and the research question.

4.2 Violations of Strict Exogeneity

Together, conditions (ii) and (iii) imply that the mechanism M is excludable. More formally, the strict exogeneity of M implies that $P(U|M,X) = P(U|X)$ and $P(Y|X,M,U) = P(Y|M,U)$. In this sub-section, we examine violations of this assumption. Again, we do this with a simulation analysis.

Our simulation setup is the same as in section 3, except that here we allow for the endogeneity of M. Let $U_i \sim N(0,1)$, $Z_i \sim U(0,1)$, $\epsilon_{Xi} \sim N(0,1)$, $\epsilon_{Mi} \sim N(0,1)$, and $\epsilon_{Yi} \sim N(0,1)$ for a sample size of $N = 100,000$ observations. Then, let

$$X_i = 0.5U_i + \epsilon_{Xi}, \tag{28}$$

$$M_i = Z_i X_i + \Gamma U_i + \epsilon_{Mi}, \tag{29}$$

and

$$Y_i = 0.5M_i + 0.5U_i + \epsilon_{Yi}. \tag{30}$$

The critical difference here is that now, when defining M in equation 29, U is included on

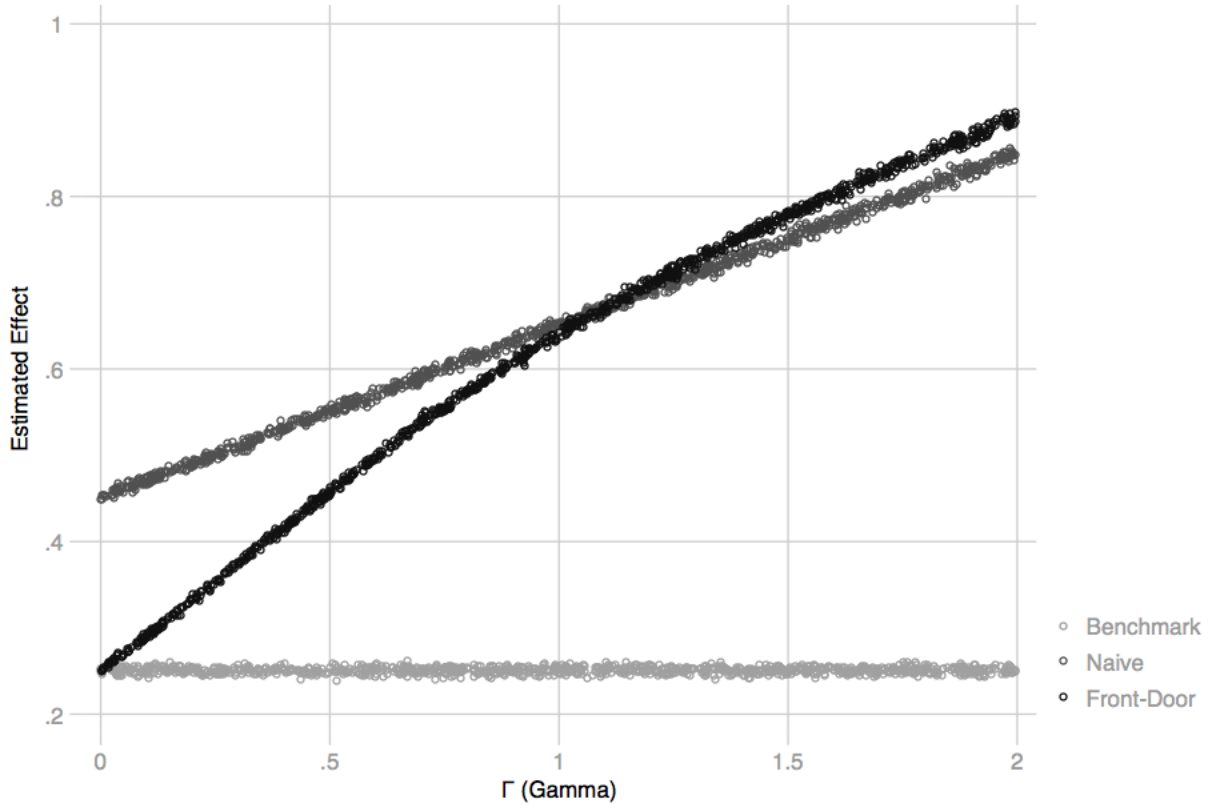
the right-hand-side. The parameter Γ defines the strength of the relationship between U and M. In this simulation analysis we let $\Gamma \sim U(0, 2)$. By permitting the values of Γ to vary allows the degree of endogeneity in our simulations to vary.

We show these simulation results graphically. Figure VI illustrates how having an endogenous mechanism influences the credibility of using the FDC approach to estimate the ATE. This figure shows estimated effects for three estimation approaches. First, the benchmark estimates (light gray), which include the confounder U on the right-hand-side of the regression equation, accurately estimates the true ATE of 0.250. Second, the naïve estimates (dark gray), which omits the confounder U from the regression equation, consistently overestimates the ATE. The size of this bias is increasing in the strength of the endogeneity of M. This is because, as Γ increases, the influence of the confounder U in the relationship between X and Y increases. Finally, the front-door estimates (black) are estimated as described in equations 18 and 19 in section 2.

Once again, a few remarks are in order. First, and rather unsurprisingly, it is only when the degree of endogeneity of M is negligible (i.e., when Γ is infinitesimally close to zero) that the FDC approach accurately estimates the ATE. Second, when M is weakly endogenous (i.e., when $\Gamma > 0$ but still relatively small) the FDC approach produces biased estimates of the ATE, but these estimates are less biased than the naïve estimates. Third, when M is strongly endogenous the FDC approach produces estimates of the ATE that are worse—that is, more biased—than the naïve estimates.

These details lead to an important discussion for applied researchers who may want to implement the FDC approach in their given empirical setting. In many cases, strict exogeneity of M may be debatable. Indeed, outside of an experimental setting, convincingly arguing that $P(U|M, X) = P(U|X)$ and $P(Y|X, M, U) = P(Y|M, U)$ will likely be challenging. That said, however, if applied researchers can convincingly argue that the degree of endogeneity of M is relatively weak—that M is not strictly exogenous but that it is plausibly exogenous (Conley et al., 2012), so to speak—then the FDC approach will

FIGURE VI: The Consequences of an Endogenous Mechanism



Notes: This figure illustrates simulation results using 1,000 replications from each estimation approach. The vertical axis represents the estimated effect. The horizontal axis represents the Gamma parameter, representing the degree of endogeneity, from equation 29. The benchmark estimates (light gray) all accurately estimate the true ATE of 0.250. The naive estimates are shown in dark gray and the front-door estimates are shown in black.

produce more reliable estimates of the ATE compared to the naïve approach which consists in regressing Y on an endogenous X. On the other hand, when the endogeneity of M is obviously relatively strong, using the FDC approach could lead to more bias in estimates of the ATE than the naïve approach. Specifically in our simulation set-up, the FDC estimates begin to become just as biased as the naïve estimates when Γ is equal to one. In the way we've defined our variables, this means that the direct effect of U on M is about twice as strong as the indirect effect of U on M via X. Of course when using real-world data, when we cannot observe U, testing the specific size of these relationships is impossible. In nearly all practical settings, the case for the exogeneity of M will rely on careful reasoning based on the given empirical setting.

4.3 Treatment Totally Defined by Mechanism

Recall that in addition to the three assumptions in section 2.1 for the FDC to identify the average treatment effect, Pearl (2000) makes a fourth assumption, namely that $P(X_i|M_i) > 0$. This assumption implies that for every value of the mechanism M, the likelihood that an observation will receive treatment X is nonzero. In other words, the treatment cannot be totally defined by the mechanism, and no matter what value the mechanism takes, it has to be the case that an observation has a nonzero probability of receiving the treatment.

This assumption was satisfied in our simulation results, but in both of our real-world examples, it failed to hold. In the Beaman et al. (2013) example, a farmer who received no fertilizer had a likelihood of receiving treatment equal to zero, as they had been assigned to the control group. Likewise, in the Bellemare et al. (2019) example, a subject who faced a standard deviation of the price distribution of zero had a likelihood of experiencing price risk equal to zero, as that subject had been assigned to a price certainty treatment.

Though our two real-world examples come from experiments, cases where $P(X_i|M_i) = 0$ are not the exclusive preserve of experimental research designs, and it is not difficult to imagine observational research designs where only those subjects who select into receiv-

TABLE V: Over-Controlling for the Treatment — Fertilizer Use in Mali

	Use of Fertilizer		Fertilizer Quantity		Fertilizer Expenses	
	(1)	(2)	(3)	(4)	(5)	(6)
Benchmark Front-Door ATE	0.603***		26.64***		-2,605.3***	
	(0.030)		(3.002)		(389.7)	
Over-Controlled Front-Door ATE		0.009		17.580***		-882.72
		(0.055)		(5.920)		(776.90)
Observations	378	378	378	378	377	377

Notes: Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All columns include the same control variables as in Beaman et al. (2013). Columns (1), (3), and (5) represent benchmark seemingly unrelated regression FDC estimates of the ATEs of receiving fertilizer on the outcome variables. Columns (2), (4), and (6) represent seemingly unrelated regression estimates of the front-door criterion ATEs which over-control for treatment in the outcome regression of the FDC setup. Standard errors are estimated by the delta method.

ing a given treatment can actually receive that treatment in nonzero amounts.

Preliminary empirical work for this paper, however, uncovered the following fact: It is only when there are no unobserved confounders that $P(X_i|M_i) = 0$ is a problem. In such cases, one only need to omit the treatment variable X from estimation in Equation 8 for the method we outline in section 2 to recover the correct ATE. When there are unobserved confounders, the method in section 2 applies lock, stock, and barrel.

Why does the treatment need to be omitted from Equation 8 when the assumption that $P(X_i|M_i) > 0$ is violated and there are no unobserved confounders? In such cases, the variation in X is already accounted for in the variation in M. Indeed, when $M_i > 0$, we know $X_i = 1$, and when $M_i = 0$, we know $X_i = 0$. Although it is certainly possible to estimate Equation 8 when the treatment is totally defined by the mechanism, both with and without unobserved confounders, it is only in the former case that the inclusion of both M and X as regressors on the left-hand side of Equation 8 will return an unbiased ATE.

The results in Table V revisit the data in Table II. Here, however, odd-numbered columns report the correct ATEs (i.e., ATEs obtained as in Equations 9, but omitting the treatment variable in Equation 8), and even-numbered columns report biased ATEs (i.e., ATEs obtained as in Equations 9, but including the treatment variable in Equation 8).

Fortunately, the problem discussed in this section is only a problem in cases without the same unobserved confounder U affecting both the treatment X and the outcome Y. Such cases, however, only really arise when X and M are experimentally assigned, in which case the FDC is not necessary to estimate the ATE—a simple regression of Y on X will return that ATE. As such, the fact that $P(X_i|M_i) > 0$ may not hold remains a largely theoretical issue compared to the other issues discussed in this paper.

5 Conclusion

This paper has focused on the application of Pearl's (1995, 2000) front-door criterion. Because the goal of most empirical social science is to answer questions of the form "Does X cause Y?," social scientists should welcome the addition of techniques that allow answering such questions to their empirical toolkit. Yet the adoption of the front-door criterion has so far been slow, with the exception of two recent articles by Glynn and Kashin (2017, 2018).

We focus here first and foremost on explaining how to use the front-door criterion in the context of linear regression, which remains the workhorse of empirical social science. We then present three empirical examples: one using simulated data, one relying on field-experimental data by Beaman et al. (2013), and one relying on experimental data from lab and lab-in-the-field experiments by Bellemare et al. (2019). Finally, in an effort to help overcome social scientists' resistance to incorporating the front-door criterion in their empirical toolkit, we look at what happens when the assumptions underpinning the front-door criterion are violated.

Our results lead to the following recommendations for applied work:

1. Because the FDC estimand is the product of two estimated coefficients from two separate regressions, it is best to estimate those two regressions simultaneously to account for the correlation in the errors. We do so in this paper by seemingly unre-

lated regression methods.

2. With control variables, both the mechanism and the outcome regression components of the FDC setup should include all controls.

3. Because the FDC estimand is a nonlinear combination of two estimated coefficients, standard errors can be computed either by the delta method or by bootstrapping. In small samples, bootstrapping should be preferred to the delta method.

4. When the treatment operates through more than one mechanism, the average treatment effect is the sum of the partial treatment effects (i.e., the "indirect effects"), defined by the effect of the treatment on outcome through each mechanism.

5. When the mechanism is no longer strictly exogenous, the usefulness of the FDC depends on the degree of exogeneity of the mechanism. In cases where the mechanism is only plausibly—but not strictly—exogenous (Conley et al., 2012), the estimate of the ATE obtained by the FDC is closer to the true value of the ATE than the estimate of the ATE obtained by a naïve regression of outcome on treatment. In cases where the mechanism is deemed to be strongly endogenous, the estimate of the ATE obtained by the FDC is further from the true value of the ATE than the estimate of the ATE obtained by a naïve regression of outcome on treatment.

6. In the absence of unobserved confounders, when the treatment is totally defined by the mechanism, it is necessary to exclude the treatment variable from the outcome regression component of the FDC setup because the information in the treatment variable is already contained in the mechanism variable, and including the treatment variable biases the ATE.

7. The FDC is most promising in cases where units of observations are selected into treatment on the basis of unobservables which also affect the outcome, but for which treatment intensity can argued to be (as good as) randomly assigned.

Ultimately, the front-door criterion is a useful tool for applied econometricians interested in causal inference with observational data. When selection into treatment is endogenous but there exists a single, plausibly exogenous mechanism whereby the treatment causes the outcome, the front-door criterion can be argued to credibly identify the causal effect of treatment on outcome.

References

Acharya, A., Blackwell, M., and Sen, M. (2016) "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects," *American Political Science Review*, vol. 110, no. 3, pp. 512-529.

Beaman, L., Karlan, D., Thuysbaert, B., and Udry, C. (2013) "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali," *American Economic Review: Papers & Proceedings*, vol. 103, no. 3, pp. 381-386.

Bellemare, M.F., Y.N. Lee, and D.R. Just (2019) "Producer Attitudes toward Price Risk: Experimental Evidence from the Lab and from the Field," Working Paper, University of Minnesota.

Conley, T.G., C.B. Hansen, and P.E. Rossi (2012) "Plausibly Exogenous," *Review of Economics and Statistics*, vol. 94, no. 1, pp. 260-272.

Glynn, A.N. and Kashin, K. (2017) "Front-Door Difference-in-Differences Estimators," *American Journal of Political Science*, vol. 61, no. 4, pp. 989-1002.

Glynn, A.N. and Kashin, K. (2018) "Front-Door Versus Back-Door Adjustment With Unmeasured Confounding: Bias Formulas for Front-Door and Hybrid Adjustments With Application to a Job Training Program," *Journal of the American Statistical Association*, vol. 113, no. 523, pp. 1040-1049.

Hernán, M.A. and Robins, J.M. (2019) *Causal Inference*, Unpublished manuscript.

Imai, K., Keele, L., and Yamamoto, T. (2010) "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects," *Statistical Science*, vol. 25, no. 1, pp. 51-71.

Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011) "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies," *American Political Science Review*, vol. 105, no. 4, pp. 765-789.

Imbens, G.W. (2019) "Potential Outcomes and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics," Working Paper, Stanford

University.

Pearl, J. (1995) "Causal Diagrams for Empirical Research," *Biometrika*, vol. 82, no. 4, pp. 669-688.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, United Kingdom.

Pearl, J. and Mackenzie D. (2018) *The Book of Why: The New Science of Cause and Effect*, Basic Books: New York, NY.

Sandmo, A. (1971) "On the Theory of the Competitive Firm Under Price Uncertainty," *American Economic Review*, vol. 61, no. 1, pp. 65-73.

Appendix

A1. Multiple Mechanisms–Case 2

We now show the results of a simulation that demonstrate the consequences (or lack thereof) of multiple mechanisms of the sort illustrated in Figure V, where multiple mechanisms lie on the same path from X to Y.

Our simulation setup is as follows. Let $U \sim N(0,1)$, $\epsilon_X \sim N(0,1)$, $Z_1 \sim U(0,1)$, $Z_2 \sim U(0,1)$, $\epsilon_{M1} \sim N(0,1)$, $\epsilon_{M2} \sim N(0,1)$, and $\epsilon_Y \sim N(0,1)$ for a sample size of $N = 100,000$ observations. Then, let

$$X_i = 0.5U_i + \epsilon_{Xi}, \tag{31}$$

$$M_{1i} = Z_{1i} \cdot X_i + \epsilon_{M1i}, \tag{32}$$

$$M_{2i} = Z_{2i} \cdot M_{1i} + \epsilon_{M2i}, \tag{33}$$

and

$$Y_i = 0.5M_{2i} + 0.5U_i + \epsilon_{Yi}. \tag{34}$$

As illustrated in Figure V, this fully satisfies Pearl's (1995, 2000) three criteria for the FDC to be able to estimate the average treatment effect of X on Y. By substituting Equation 31 into Equation 32, then substituting equation 32 into equation 33, and finally substituting equation 33 into Equation 34, it should be immediately obvious to the reader that the true ATE is equal to 0.125 in our simulations.

Similar to the previous simulation analysis, we estimate two specifications. The first specification estimates the true ATE by controlling for the confounder U. The second specification estimates the ATE using the FDC approach. As the results in Table V show, estimates of the ATE with the FDC approach in this case are statistically invariant whether either or both M_1 and M_2 are included in the estimation procedure.

TABLE VI: Simulation Results—Multiple Mechanisms, Case 2

	Benchmark	Naïve	Front-Door (Both)			Front-Door (M_1 only)		Front-Door (M_2 only)	
Variables	Y	Y	M_1	M_2	Y	M_1	Y	M_2	Y
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment (X)	0.127***	0.326***	0.495***	0.245***	0.201***	0.496***	0.120***	0.245***	0.202***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
Mechanism (M_1)	–	–	–	–	0.003	–	0.254***	–	–
					(0.004)		(0.004)		
Mechanism (M_2)	–	–	–	–	0.502***	–	–	–	0.503***
					(0.003)				(0.003)
Confounder (U)	0.501***	–	–	–	–	–	–	–	–
	(0.004)								
Intercept	-0.001	0.004	0.002	0.004	0.002	0.002	0.004	0.004	0.002
	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
Estimated ATE	0.127***	0.326***		0.125***		0.126***		0.123***	
	(0.004)	(0.004)		(0.002)		(0.002)		(0.002)	
Observations	100,000	100,000		100,000		100,000		100,000	

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The front-door equations in columns (3) and (4) are estimated by seemingly unrelated regressions. The standard error for the front-door ATE is estimated by the delta method.