

ERRATA CORRECTIONS FOR SECOND EDITION OF *CAUSALITY*

~~7/14/09 Updated~~

page 10 *append* to line 6 to read, “and the *regression coefficient* (of X on Y)”

page 11 *change* formula in line 5 after Eq. (1.29) to read, “ $(X \perp\!\!\!\perp Y|Z)$.”

page 24 *insert* after Eq. (1.37), “(and implies) conditions (i)–(iii) of” [continue with “which follows from...”]

page 64 *append* sentence to end second paragraph of **Postscript for the Second Edition**, “Spirtes, Glymour, Scheines, and Tillman (2010) summarize the current state of the art in causal discovery.”

page 100 second paragraph, line 1, *replace* “to some statisticians,” with “to traditional statisticians,”

page 105 last paragraph before **Applications and Critics**, *add* sentence after footnote 16 marker: “Tian and Shpitser (2010) provide a comprehension summary of these results.”

page 105 *change* in last line before footnotes 15 and 16, “(Pearl 2008).” to “(Pearl 2009a).”

page 106 end of 2nd paragraph: *change* last line to read, “criticized this attitude; Pearl (2009ab, 2010a) highlights its damaging consequences.”

page 106 *add* sentence to last line of page, “This semantics will enable us to develop a number of techniques for counterfactual analyses (Chapters 8–11), including the Mediation Formula (equations (4.17)–(4.18)) – a key tool for assessing causal pathways in nonlinear systems.”

page 132 *replace* complete page to read:

4.5.5 Indirect Effects and the Mediation Formula

Remarkably, the definition of the natural direct effect (4.11) can easily be turned around and provide an operational definition for the *indirect effect* – a concept shrouded in mystery and controversy, because it is impossible, using the $do(x)$ operator, to disable the direct link from X to Y so as to let X influence Y solely via indirect paths.

The natural indirect effect, IE , of the transition from x to x' is defined as the expected change in Y affected by holding X constant, at $X = x$, and changing Z to whatever value it would have attained had X been set to $X = x'$. Formally, this reads (Pearl 2001c):

$$IE_{x,x'}(Y) \triangleq E[(Y(x, Z(x'))) - E(Y(x))], \quad (4.14)$$

We see that, in general, the total effect TE of a transition is equal to the *difference* between the direct effect of that transition and the indirect effect of the reverse transition:

$$TE_{x,x'}(Y) \triangleq E(Y(x') - Y(x)) = DE_{x,x'}(Y) - IE_{x',x}(Y). \quad (4.15)$$

In linear models, where reversal of transitions amounts to negating the signs of their effects, (4.15) provides a formal justification for the standard additive formula

$$TE_{x,x'}(Y) = DE_{x,x'}(Y) + IE_{x,x'}(Y). \quad (4.16)$$

In the simple case of unconfounded mediators, the natural direct and indirect effects are estimable through two regression equations called the Mediation Formula:

$$DE_{x,x'}(Y) = \sum_z [E(Y|x', z) - E(Y|x, z)]P(z|x). \quad (4.17)$$

$$IE_{x,x'}(Y) = \sum_z E(Y|x, z)[P(z|x') - P(z|x)] \quad (4.18)$$

These provide two ubiquitous measures of mediation effects, applicable to any nonlinear system, any distribution, and any type of variables (Pearl 2009b, 2010b).

Note that the indirect effect has clear policy-making implications. For example: in a hiring discrimination environment, a policy maker may be interested in predicting the gender mix in the work force if gender bias is eliminated and all applicants are treated equally—say, the same way that males are currently treated. This quantity will be given by the indirect effect of gender on hiring, mediated by factors such as education and aptitude, which may be gender-dependent. See (Pearl 2001c, 2010b) for more examples.

More generally, a policy maker may be interested in the effect of motivating a select set of subordinate employees, or of controlling the routing of

messages in a network of interacting agents. Such applications motivate the analysis of *path-specific effects*, that is, the effect of X on Y through a selected set of paths (Avin et al. 2005).

In all these cases, the policy intervention invokes the selection of signals to be sensed, rather than variables to be fixed. Pearl (2001c) has suggested therefore that signal sensing is more fundamental to the notion of causation than manipulation; the latter being but a crude way of stimulating the former in experimental setup (see Section 11.4.5). A general characterization of counterfactuals that are empirically testable is given in Chapters 7, 9, 11, and in Shpitser and Pearl (2007).

page 169 line 1, *change* “dichotomous, then” to read “dichotomous and the parameter of interest is $\lambda = P(y_0|x_0)$, then”

page 169 line 4, *change* “as well as their ratio:” to read “and λ “involves” their ratio:”

page 169 *change* footnote 25 to read: “Engle et al. (1983, p. 281) and Hendry (1995, pp. 162–3) overcome this ambiguity by using selective “reparameterization” – a necessary step which textbooks tend to ignore.”

page 171 Section 5.6.1, line 12 of first paragraph, *change* “in the models.” to “in the model.”

page 172 line 6, *change* last sentence to read:
“See McDonald (2002a) for an algebraic approach, and Brito (2010) for a gentle introduction and a survey of results.”

page 172 Section 5.6.3, line 4, end of second formula, *change* “=
 $E(Y|do(z));$ ” to “=
 $E(Y|do(x));$ ”

page 240 *append* sentence to end of footnote 21, “Pearl (2010d) discusses the counterfactual interpretation of $do(A \text{ or } B)$.”

page 274 *append* sentence to last paragraph before **The Instrumental Inequality**, “Richardson and Robins (2010) discuss the power of these tests.”

page 330 *add* new paragraph to end of page:
“Halpern and Hitchcock (2010) summarize the state of the art of the structural approach to actual causation, and discuss its sensitivity to choice of variables.”

page 335 *append* to first paragraph of Section 11.1.2, “(See also Hayduk et al., 2003.)”

page 339 end of line 3 of *Author’s Answer*., the “|” in the formula $P(y|x)$ is printed too heavy.

page 339 -3 lines before bottom of page, *insert* “(or some $e_i, i > 0$)” *after* “some S_i ”

page 340 line 16–17, *replace* “for any consequence of treatment (not only” with “for consequences of treatment that open such spurious paths (not only”

page 344 line 12, *replace* “Chapter 9).” with “Chapter 9 and Shpitser and Pearl 2007).”

page 346 paragraph 3, line 7: **replace** section starting, “For example, $\{Z_1, V\}$, $\{Z_2, V\}$, or $\{Z_1, Z_2\}$, ...”

with

For example, V and Z_2 can be removed from C by successively applying conditions C_1 and C_2 , thus producing an irreducible subset, $\{Z_1, W_1, W_2\}$, c -equivalent to the original covariate set C . However, this subset is inadmissible for adjustment because, like C , it does not satisfy the back-door criterion.

continue with “While a theorem ...”

page 346 **replace** first equation of Section 11.3.4 **with**:

$$\begin{array}{ll} V \perp\!\!\!\perp \{W_1, W_2\} & X \perp\!\!\!\perp \{V, Z_2\} | \{Z_1, W_2, W_1\} \\ Z_1 \perp\!\!\!\perp \{W_2, Z_2\} | \{V, W_1\} & V \perp\!\!\!\perp Y | \{X, Z_2, W_2, Z_1, W_1\} \\ Z_2 \perp\!\!\!\perp \{W_1, Z_1, X\} | \{V, W_2\} & V \perp\!\!\!\perp Y | \{Z_2, W_2, Z_1, W_1\} \end{array}$$

page 347 correction to figure caption 11.9

page 346-7 **Replace** text from

“A less trivial example...”

through

“...not possibly have direct effect on Y .”

with

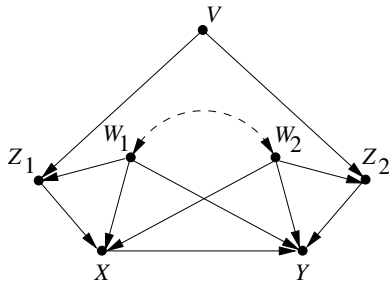


Figure 11.9: A model that is almost indistinguishable from that of Figure 11.8(b), save for advertising one additional independency $Z_1 \perp\!\!\!\perp Y | X, W_1, W_2, Z_2$. It deems three sets to be admissible (hence c -equivalent): $\{V, W_1, W_2\}$, $\{Z_1, W_1, W_2\}$, and $\{W_1, W_2, Z_2\}$, and would be rejected therefore if any pair of them fails the c -equivalence test.

A less trivial example, one that is not sensitive to choice of parameters, lies in the class of equivalent structures, in which all conditional independencies emanate from graph separations. The search techniques developed in Chapter 2 provide systematic ways of representing all equivalent models compatible with a given set of conditional independence relations.

The model depicted in Figure 11.9 is a tough contender to that of Figure 11.8(b); it satisfies all the conditional independencies implied by the latter, plus one more: $Z_1 \perp\!\!\!\perp Y | X, W_1, W_2, Z_2$, which is not easy to detect or test. Yet, contrary to Figure 11.8(b), it deems three sets $\{Z_1, W_1, W_2\}$, $\{V, W_1, W_2\}$, and $\{Z_2, W_1, W_2\}$ to be admissible, hence c -equivalent; testing for the c -equivalence of the three sets should decide between the two contesting models.

Substantive causal knowledge may provide valuable information for such decisions. For example, the model of Figure 11.9 can be ruled out if we have good reasons to believe that variable W_2 cannot have any influence on X (e.g., it may occur *later* than X), or that W_1 could not possibly have direct effect on Y .

page 350 last paragraph, line 6: *change* “in the restricted language” to “in the cryptic language”

page 351 second paragraph, line 7: *delete* “not only is not needed, but”

page 351 second paragraph, line 8: *add* “, Pearl 2010??” after “Sjölander 2009”

which pearl:10?

page 356 *replace* second paragraph of **Author’s Reply**: with:

This counterfactual definition leads to the Mediation Formula (4.18) which extends path analysis to nonlinear models and enables us to estimate indirect effects for categorical variables using ordinary regression.

page 358 second line: *change* second line to read, “and Greenland (1992), and to the Mediation Formula of equations (4.12), (4.17), and (4.18).” [continue with “Enlightened by these...”]

page 358 in lines 7-8: *change* to read, “Pearl (2001), Petersen et al. (2006), and Pearl (2010c).”

page 358 line 4 of **Author’s Reply**:, *change* “alone is manipulated,” to “alone is manipulated directly,”

page 358 second paragraph, line 3 of **Author’s Reply**:, *change* “was far from ideal,” to “was far from the one intended,”

page 358 third paragraph, line 2 of **Author’s Reply**:, *change* “model are affected” to “model are directly affected”

page 358 third paragraph, line 3 of **Author’s Reply**:, *change* “specify explicitly and formally” to “specify formally”

page 358 *append* sentence to last line of **Author’s Reply**:, , “See Section 11.4.6.”

page 359 in **Author’s Answer**: *replace* “of “intention to treat.”)” with “of “intention to treat,” and p. 363 for an example.)

page 360 line 2–3, *replace* sentence starting “LISREL’s formula...” to read: Shpitser and Pearl (2009) provide a sufficient and necessary graphical condition for identifying the effect of the “add n to X ” operator.

page 366 *replace* the last two paragraphs of 11.5.1 with one paragraph: “The intuition is that, conditional on Z , W acts as an instrumental variable relative to $X \rightarrow Y$. See also McDonald (2002a). More general identification methods are reported in Brito and Pearl (2002a,b,c; 2006), and surveyed in Brito (2010).”

page 391 *replace* line 1-2 of **Conjecture**: with,

Any counterfactual query of the form $P(Y_x|e)$ is empirically identifiable in every constant-effect model, i.e., $Y_{x_1}(u) - Y_{x_2}(u)$ is constant over u .

Corrections to Subject Index

- page 461** *change* under subject heading “direct effects”, sub-subject “average (natural)...” to read “natural, 130–1, 355–8” (and move after “identification (parametric)...” also *add* page to sub-subject “identification (nonparametric), 128, 131”
- page 461** *add* sub-subject heading, “completeness, 105” under subject heading “do-calculus”
- page 461** *add* pages to “ETT (effect of treatment on the treated)” to read: “269–70, 243–4, 390, 396”
- page 462** imaging, 73, 112, 242–3, 461
- page 463** *add* subject heading “MGM” before “Markov” subject heading
- page 463** *add* sub-subject heading under “mechanisms, 22”:
“mediation *see* indirect effects”
- page 463** *add* subject “Mediation Formula, 106¹³², 358”
- page 463** *add* subject “Neyman-Rubin model” after “Newcomb’s paradox...” then *add* sub-subject “*see* Potential outcome framework”
- page 359** *change* “effect of “intention to tre

[Avin et al., 2005] C. Avin, I. Shpitser, and J. Pearl. Identifiability of path-specific effects. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence IJCAI-05*, pages 357–363, Edinburgh, UK, 2005. Morgan-Kaufmann Publishers.

[Brito, 2010] C. Brito. Instrumental sets. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 295–???. College Publications, London, 2010.

[Halpern and Hitchcock, 2010] J.Y. Halpern and C. Hitchcock. Actual causation and the art of modeling. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 383–???. College Publications, London, 2010.

[Pearl, 2001] J. Pearl.at.”)” to “effect of “intention to treat,” and p. 363 for an example.)”

~~360~~

Corrections to Name Index

- page 454** Dawid, A.P., *remove* “132,”
- page 456** Richard, T.S., *insert* “274”
- page 456** Robins, J.M., *insert* “274”

page 457 Shpitser, I., *insert* “344”

2.23.10

page 337 - heading ”Typical application:” and ”Remark on correlated errors:” one is bold the other is bold italic

check page 346 vs file (ch11-3-revised): For example, V and Z_2 can be removed from C by successively applying conditions

page 346: ‘none of these subsets is inadmissible’ ‘none of these subsets are inadmissible’

add new cites to refs:

pearl 2010a = [Pearl, 2010c]

this is actually 2009, not 2010

pearl 2010b = [Pearl, 2009]

pearl 2010c = [Pearl, 2010a]

pearl 2010d = [Pearl, 2010b]

pearl:09-??

[Brito, 2010]

[Halpern and Hitchcock, 2010]

need cite for hayduk:etal03

[?]

[Richardson and Robins, 2010]

[Shpitser and Pearl, 2009]

[Spirtes *et al.*, 2010]

[Tian and Shpitser, 2010] tribute book, pp 523–544

Bibliography

- [Avin *et al.*, 2005] C. Avin, I. Shpitser, and J. Pearl. Identifiability of path-specific effects. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence IJCAI-05*, pages 357–363, Edinburgh, UK, 2005. Morgan-Kaufmann Publishers.
- [Brito, 2010] C. Brito. Instrumental sets. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 295–???. College Publications, London, 2010.
- [Halpern and Hitchcock, 2010] J.Y. Halpern and C. Hitchcock. Actual causation and the art of modeling. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 383–???. College Publications, London, 2010.
- [Pearl, 2001] J. Pearl. Direct and indirect effects. In *Uncertainty in Artificial Intelligence, Proceedings of the Seventeenth Conference*, pages 411–420. Morgan Kaufmann, San Francisco, CA, 2001.
- [Pearl, 2009] J. Pearl. Causal inference in statistics: An overview. *Statistics Surveys*, 3:96–146, <<http://www.bepress.com/ijb/vol6/iss2/7/>>, 2009.
- [Pearl, 2010a] J. Pearl. The mediation formula: A guide to the assessment of causal pathways in non-linear models. Technical Report R-363, <http://ftp.cs.ucla.edu/pub/stat_ser/r363.pdf>, Department of Computer Science, University of California, Los Angeles, CA, 2010.
- [Pearl, 2010b] J. Pearl. Physical and metaphysical counterfactuals. Technical Report R-359, <http://ftp.cs.ucla.edu/pub/stat_ser/r359.pdf>, Department of Computer Science, University of California, Los Angeles, CA, 2010. Submitted to *Review of Symbolic Logic*.
- [Pearl, 2010c] J. Pearl. The structural theory of causation. In F. Russo P. McKay Illari and J. Williamson, editors, *Causality in the Sciences*. Clarendon Press, Oxford, 2010. Forthcoming.
- [Richardson and Robins, 2010] T.S. Richardson and J. Robins. Analysis of the binary instrumental variable model. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 415–???. College Publications, London, 2010.

- [Shpitser and Pearl, 2007] I. Shpitser and J Pearl. What counterfactuals can be tested. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*, pages 352–359. AUAI Press, Vancouver, BC, Canada, 2007. Also, *Journal of Machine Learning Research*, 9:1941–1979, 2008.
- [Shpitser and Pearl, 2009] I. Shpitser and J Pearl. Effects of treatment on the treated: Identification and generalization. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, Montreal, Quebec, 2009.
- [Spirtes *et al.*, 2010] P. Spirtes, C. Glymour, R. Scheines, and R Tillman. Automated search for causal relations: Theory and practice. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 467–???. College Publications, London, 2010.
- [Tian and Shpitser, 2010] J. Tian and I. Shpitser. On identifying causal effects. In R. Dechter, H. Geffner, and J.Y. Halpern, editors, *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pages 523–???. College Publications, London, 2010.