Coasean Causal Inference: Using Models as Foils

Brian C. Albrecht and Brian Kogelmann

Ronald Coase reasoned with economic models, and made inferences more generally, in an unorthodox manner. In "The Problem of Social Cost", Coase draws a conclusion concerning the impact of transaction costs in the real world by looking to a model world in which transaction costs are conspicuously absent. Yet how Coase draws such an inference – from the model world to the actual world – remains opaque. The central goal of this paper is to make this method of inference crystal clear. We achieve this by offering a formal characterization of Coase's method of inference; we call this form of reasoning *models as foils*. We then show that other important thinkers from the history of economic thought also used models as foils, and we end by arguing that this method of inference can be used to fruitful ends by contemporary economists. Fresh insights can be gained if one interprets existing models in a Coasean manner.

1. Introduction

Ronald Coase stands as one of the most important economists of the 20th century. Though his greatness is now taken for granted, it was by no means a matter of course that he would secure such an illustrious place in the economics profession. For though his contributions are now etched into the mainstream of the discipline, Coase himself pursued his work in deeply heterodox fashion. At a time when economics was becoming a more technical discipline, Coase eschewed the use of mathematics, famously jesting that: "In my youth it was said that what was too silly to be said may be sung. In modern economics it may be put into mathematics" (Coase 1988a: 185). And at a time when economics was becoming more insulated from other fields of academic inquiry – transitioning from *political economy* to just *economics* – he relied on history and law to draw important conclusions in economic theory.[1]

The divergence between Coase and mainstream economics does not end with his eschewal of mathematics or embrace of law and history. Little discussed in the massive literature on Coase

[1] For a broad overview of Coase's approach to economics, see Medema (1995).

is the unorthodox way Coase reasons with economic models and draws inferences more generally. In his most influential paper "The Problem of Social Cost," Coase draws a conclusion concerning the impact of transaction costs in the real world by looking to a model world in which transaction costs are conspicuously absent. Yet how Coase draws such an inference – from the model world to the actual world – remains opaque. The central goal of this paper is to make this method of inference crystal clear. We achieve this goal by offering a formal characterization of Coase's method of inference.

There are important payoffs to doing this. By offering a formal characterization of Coase's reasoning, we can see that other important figures from the economics profession embrace a similar methodology, although this similarity is often ignored. More specifically, we show that the method of inference Coase uses also describes how several important thinkers from the Austrian school of economics reason with models. This helps confirm Peter Boettke's thesis that there exists a deep similarity between Coase and the Austrians (e.g., Boettke 1997; Boettke 2012). Not only this, but we believe Coase's method of inference can be used to fruitful ends by contemporary economists as a new way of approaching what is now a decisively model-based discipline. Fresh insights can be gained if one interprets existing models in a Coasean manner.

Our paper thus makes contributions to three separate literatures. First, the paper adds to the literature on the methodology of Coase, by offering a formal characterization of Coase's method of inference.[2] Second, the paper contributes to the history of political economy, by drawing methodological connections between thinkers that initially do not seem to have much in common. And third, the paper contributes to the literature on economic methodology more generally, as it proposes a new way in which economists can reason with models.[3]

The structure of this paper is as follows. The next section examines "The Problem of Social Cost," where Coase famously examines a model world in which transaction costs are absent. The point of examining this model is not to derive theorems about the properties of such an unrealistic world (though this has been the takeaway of some), but rather to derive insights about our actual world. How Coase jumps from the model world to the real world, though, remains opaque. Section three clears this up by offering a formal characterization of Coase's method of inference. Inspired

[2] Recent work on Coase's methodology includes Allen (2015); Bertrand (2016).
[3] For a broad overview of recent work on economic methodology, see Mireles-Flores (2018).

by John Stuart Mill's work on inductive inferences, we call this method of inference *models as foils*. With a better understanding of Coase's method of inference under our belts, section four shows that models as foils accurately characterizes the reasoning used by other important and influential economists – in particular, those of the Austrian school. We end in section five by suggesting how Coase's methodology can still help us today. In particular, we show that fresh insights can be gleaned by treating models in the contemporary literature as foils.

2. Coase on Transaction Costs

Coase begins his 1960 paper by telling us that he is interested in "those actions of business firms which have harmful effects on others" (Coase 1960/1988: 95). That is, he is interested in externalities: costs which parties not privy to voluntary exchange pay nonetheless. It is intuitively plausible that underlying property law matters for how externalities get resolved. If Althea wants to run loud machinery on her property, and Bertha wants peace and quiet on her adjacent property, then how this dispute will be resolved seems to depend on the allocation of rights: does Althea have a right to run machinery, or does Bertha have a right to peace and quiet? If Althea has the right, then the machinery will run; if Bertha has the right, then it will not.

But, Coase shows, this is not necessarily so. Coase begins his analysis by assuming that "the pricing system works smoothly (strictly this means that the operation of a pricing system is without cost)" (Coase 1960/1988: 97). That is, he assumes there are zero transaction costs. From here Coase goes through examples showing that, in such a world, (*i*) all disputes are resolved in an efficient manner, but more stunningly (*ii*) all disputes result in the *same* efficient outcome regardless of the initial allocation of property rights. In other words, the initial allocation of property rights is irrelevant in the frictionless world of zero transaction costs. He writes: "with costless market transactions, the decision of the courts concerning liability for damage would be without effect on the allocation of resources" (Coase 1960/1988: 106). To continue our toy example from the paragraph above, if there are no transaction costs then the dispute between Althea and Bertha will (*i*) be resolved in an efficient manner and (*ii*) be resolved in the *same* manner regardless of whether Althea has a right to use noisy machinery or whether Bertha has a right to enjoy peace and quiet. In such a world, property rights are simply irrelevant.

Though it is not a mathematical model of the kind most contemporary economists build, Coase's imaginary world can be studied as an informal model, nonetheless. But, given the "very unrealistic" (Coase 1960/1988: 114) assumption of no transaction costs, what is the point of such a model? What purpose does it serve in economic inquiry? This is a controversial question. For most economists, the takeaway from this model is the so-called Coase Theorem – coined by George Stigler – which says that "under perfect competition private and social costs will be equal" (Stigler 1966: 113).[4] Coase's model is thus interpreted as another theorem of welfare economics. Absent transaction costs (the typical assumption of most general equilibrium models at the time) all externalities will be bargained to the same efficient point.

Others note that another function of "The Problem of Social Cost" is to show the redundancy of Pigouvian remedies to externalities. According to Pigou, externalities require taxation, subsidization, or regulation, lest we be left with inefficiency. Crucially, neoclassical economists who embrace the Pigouvian approach to externalities often *also* theorize about the frictionless world of zero transaction costs. Thus, Coase's work shows that Pigouvian remedies are redundant, because there would be no inefficiencies resulting from externalities in such a world. As Steven G. Medema articulates it: "What Coase intended with his analysis of the zero transaction costs world was to show that in such a world, Pigouvian remedies (e.g., taxes, subsidies, and regulations) are not necessary to resolve externality problems" (Medema 1994: 209).[5]

Beyond criticizing Pigou, the Coasean world of zero transaction costs is supposed to teach us something important about our actual world, where transaction costs are not only positive but significantly so.[6] As Elodie Bertrand notes, Coase "studied the zero transaction costs world to emphasise the influence of these costs on the result of the negotiations and on the choice of the best policy" (Bertrand 2010: 976). In his own words, Coase tells us that his goal "was *not* to describe what life would be like in such a world but to provide a simple setting in which to develop the analysis and, what was even more important, *to make clear the fundamental role which transaction costs do, and should, play in fashioning the institutions which make up the economic*

[4] For detailed discussion of how Coase's work has been interpreted and received by the economics profession, see Medema (2011a); Medema (2011b); Medema (2015); Frischmann and Marciano (2015).
[5] For further examination of this point, see Aslanbeigui and Medema (1998).
[6] The classic estimate by John Wallis and Douglass North (1986) puts transaction costs sectors of the economy at 55% of GNP in 1970.

system" (Coase 1988b: 13) (emphasis ours).[7] In other words, the point of examining a model world without transaction costs is to better understand the role transaction costs play in our actual world.

What does Coase think his model teaches us about the actual world? He writes:

> Of course, if market transactions were costless, all that matters (questions of equity apart) is that the rights of the various parties should be well defined and the results of legal action easy to forecast. But as we have seen, the situation is quite different when market transactions are so costly as to make it difficult to change the arrangement of rights established by the law. In such cases, the courts directly influence economic activity. It would, therefore, seem desirable that the courts should understand the economic consequences of their decisions and should, insofar as this is possible without creating too much uncertainty about the legal position itself, take these consequences into account when making their decisions. Even when it is possible to change the legal delimitation of rights through market transactions, it is obviously desirable to reduce the need for such transactions and thus reduce the employment of resources in carrying them out (Coase 1960/1988: 119).

We interpret this passage as Coase drawing two key conclusions from his informal model. First: it is *precisely because* of transaction costs that different allocations of property rights result in diverging allocations of resources with differing levels of efficiency in the real world. That is, when and because transaction costs are present, how the dispute between Althea and Bertha gets resolved will depend on who has the relevant right. Indeed, there is no reason to think different ways of specifying the relevant rights will lead to an efficient outcome in either case. As Deidre McCloskey summarizes: "Coase's actual point, the core of a Coasean economics, was to note what happens in the many important cases in which transaction costs can*not* be neglected. If the situation *does* have high transaction costs, then it *does* matter where the liability for pollution is placed" (McCloskey 1998: 368).

This leads to Coase's second conclusion, which is normative. Because transaction costs influence how different allocations of property rights affect economic efficiency, judges should

[7] For similar remarks about how Coase believed his model was misunderstood, see Coase (1988a: 174); Coase (1988b: 15).

take the economic impact of their decisions into account when allocating property rights.[8] We will not be interested in Coase's normative conclusion for the rest of the paper. But we will be interested in his positive conclusion. More specifically, we will be interested in how he draws this conclusion – a conclusion about the relationship between property law, allocations of resources, and efficiency *in the real world* – from a model world that is radically and purposefully unrealistic in its assumptions.

3. Models as Foils

Coase examines a model world of zero transaction costs to gain insight about the impact of transaction costs in our actual world, where transaction costs are positive and significantly so. Yet, how he draws conclusions about the real world by focusing primarily on a fantasy world is unclear.[9] We now seek to fill this lacuna. To begin, it is helpful to go back in time. In particular, we would like to briefly examine a method of inference that John Stuart Mill articulates, called the *method of difference*. Mill's method of difference and Coase's inferential strategy are not the same. However, the two are quite similar. Examining the similarities and differences between these two inferential strategies will allow us to precisely pin down an exact characterization of Coase's method of inference.

Mill describes the method of difference as follows:

> If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common scarce one, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or the cause, or an indispensable part of the cause, of the phenomenon (Mill 1843: 452).

The above was a bit abstract, but Mill offers an example that clears things up. Suppose, in Case 1, what Mill calls "elements" A B C produce an outcome characterized by properties *a b c*. Then, suppose in Case 2 that elements B C produce an outcome characterized by properties *b c*. Clearly, Case 1 and Case 2 have different outcomes. In Case 1, an outcome characterized by properties *a b*

[8] For examination of this normative conclusion, see Bertrand (2015).
[9] One exception here is Hsiung (2004). Though we do not have the space to discuss this account of Coase's analytical method, we believe that the account we offer is quite distinct from Hsiung's.

c occurs, while in Case 2 an outcome characterized by properties *b c* occurs. Since the only difference in causal elements between the two cases is the presence of A, there is reason to think that element A causes outcome *a*. This is the method of difference.

Trying to offer a bit more clarity, we believe the following sequence of reasoning is an accurate characterization of Mill's method of difference.

1. In Case 1, features $f_1, f_2, \ldots, f_n$ produce outcome *O*.
2. In Case 2, features $f_1, f_2, \ldots, f_{n-1}$ produce outcome *O*`.
3. Properties *o** are present in outcome *O*, but not present in outcome *O*`.
4. Since the only difference between Case 1 and Case 2 is feature f_n, there is reason to think that f_n causes *o**.

Stated like this, we believe that Mill's method of difference is clear enough, and certainly a reasonable way of drawing inductive inferences. The question is whether this strategy of inference mirrors Coase's reasoning with his model of a world absent transaction costs.

Initially one might think that this does capture Coase's reasoning. For consider:

1. In Case 1 (the real world), different allocations of property rights result in externalities being resolved in diverging ways (some or all of which may be inefficient).
2. In Case 2 (the real world *minus* transaction costs), all externalities are bargained to the same efficient point, regardless of the initial allocation of property rights.
3. In Case 1, we see different allocations of property rights affect the resolution of externalities. This is not so in Case 2.
4. Since the only difference between Case 1 and Case 2 is the presence of transaction costs, there is reason to think that transaction costs influence how different allocations of property rights resolve externalities.

From the conclusion stated in (4) would then follow Coase's normative corollary: that property rights should be structured to minimize inefficiencies. While this application of Mill's method of difference reaches the same conclusion that Coase does, there are nonetheless significant problems with it.

In the above schematization, step (2) says that Case 2 consists of all those features of the real world, minus transaction costs. The outcome of this case is that all externalities are bargained to the same efficient point, regardless of the initial allocation of property rights. Case 2 in the above schematization represents Coase's model world. It is true, we know, that the outcome of this model world is indeed invariance in the resolution of externalities with respect to initial allocations of property rights. The problem, though, is that Coase's model world does not consist of *all* those

features of the real world, *except* for transaction costs. In fact, there are many features of the real world that Coase's model leaves out. This should not be surprising. By definition, models abstract away from almost all features of reality in their effort to simplify.

Coase's informal model, for example, leaves out the fact that the sky is blue. More relevant than this, the model leaves out the fact that persons sometimes behave irrationally, that they are often driven by prejudice and bias, that some people suffer loss aversion, and so on. As such, it is misleading to say that Case 2 consists of all those features of the real world *except* transaction costs. Coase's model abstracts away from many features of the real world, transaction costs being but one. This fact, though, means that Mill's method of difference is inapplicable to Coase's reasoning. According to Mill's method of difference, we can only draw a causal inference between two cases so long as these two cases "have every circumstance in common scarce one" (Mill 1843: 452). Problematically, Coase's model and the real world do not have every circumstance in common scarce one. There are many features across which they differ, transaction costs being one of many. We thus cannot pin the causation down on one feature. Perhaps it is the absence of irrationality among the bargaining parties in the model world that explains the difference between the two cases.

Another way of thinking about the problem here is that Coase wants to use something like Mill's method of difference, but the two cases he wants to use are (*i*) what we observe in the real world, and (*ii*) what we observe in a model world. Yet, a model world will never be able to include all those features present in the real world "scarce one." Because of this, one might conclude that not only does Mill's method of difference fail to capture Coase's reasoning, but that it will fail to capture *any* reasoning that employs a comparison between two distinct cases, where one of the cases consists of real-world observation and the other a model world. Our account of models as foils articulated below not only captures Coase's reasoning, but allows for these sorts of comparisons more generally.

The method of difference does not capture Coase's method of inference, but it does seem quite close. In particular, both Coase's inferential strategy and the method of difference (*i*) examine two cases, (*ii*) these two cases have different outcomes, and (*iii*) these different outcomes are pinned on a particular feature that causes the relevant difference. For Mill, it is easy to pin down the relevant feature. By hypothesis, there is only one difference between the two cases, so the

differentiating feature is likely to be the causal one. Coase still wants to pin down the difference between the two cases on the presence or absence of one particular feature. But he cannot do this as easily as Mill wishes to with the method of difference, because there are many features that the real world and model world differ across, not just one.

This does not mean, however, that we cannot reasonably infer that one feature does in fact cause the relevant difference. Consider a toy example. Suppose features $\{f_1, f_2, f_3, f_4\}$ cause some outcome *O*, and features $\{f_1, f_2\}$ cause *O*`, where *o** represent those properties present in outcome *O* but absent in outcome *O*`. Here, Mill's method of difference is inapplicable because the two cases differ across multiple features – f_3 and f_4. As such, it is not immediately obvious whether f_3 is the primary cause of *o**, whether f_4 is, or whether f_3 and f_4 work together in causing *o**. But just because this is not immediately obvious, it does not follow that a single causal relation cannot be established with some degree of credibility. We know, at least, that either f_3 or f_4 are the likely culprits causing *o**. In these sorts of cases, it is sometimes (but not always) possible to offer an argument as to why it is f_3 over f_4 (or vice versa) that explains *o**. Perhaps there is a theoretically plausible mechanism showing how f_3 induces *o**, yet there is no such mechanism for f_4. Or perhaps in reasonably similar circumstances, it is f_3 that produces the outcome, not f_4. Should such an argument be established, then we do have *some* reason to insist that f_3 is the primary cause of *o**, rather than f_4, *even if* the justification of our belief here is not as strong as would be the case if the only difference between the two cases was f_3 (in which case we would then be allowed to use Mill's method of difference).

Coase does something like this. He does not simply present two cases – the real world where property rights matter in the resolution of externalities, and the model world where they do not – and then conclude that transaction costs are the relevant causal feature. Rather, he carefully works through examples in the model world where property rights are allocated differently, and argues that it is precisely because transaction costs are absent that these different allocations of rights are irrelevant. He does this by examining cases of conflict (the rancher and the farmer, the doctor and the confectioner, etc.), showing that there exist efficient bargains the parties *could* reach, *if* they had limitless time to acquire more information, engage in bargaining, and so on. And, he takes the time to show that they would reach the same bargain regardless of which party has the relevant right. We believe that carefully working through these cases and pointing out the existence

of efficient bargains that could be reached constitutes an argument for why it is the absence of transaction costs in the model world that is driving the difference in outcomes, not some other absent feature.

So Coase uses something like Mill's method of difference, only (*i*) there is more than one difference between the two cases and, because of this, (*ii*) some kind of argument needs to be given concerning why one differentiating feature (rather than any other differentiating feature) is responsible for the difference in outcomes. A bit more precisely:

1. In Case 1 (the real world), the set of features *F* produces outcome *O*.
2. In Case 2 (the model world), the set of features *F* `, which is a proper subset on *F*, produces outcome *O*`.
3. Properties *o** are present in outcome *O*, but not present in outcome *O*`.
4. Letting *F** be the complement of *F* ` (indexed to the universe *F*), and letting f_n be an element of *F**, there exists an argument for why f_n, compared to all the other elements in *F**, best explains the presence of *o**.
5. As such, there is some reason to believe that f_n causes *o**.

Call this inferential strategy *models as foils*, for a model world is being used as a foil against reality to gain insight on the causal impact of some feature of the real world. In particular, we gain insight into feature *f* of the real world by comparing the real world (step one) to the model world where *f* is not present (step two). Noting that there is a difference between the real world and the model world (step three), we argue that *f* is the most plausible explanation for those features present in the real world but absent in the model world (step four). This allows us to draw our conclusion concerning the causal efficacy of real-world feature *f*.

Some things to note about models as foils. Models as foils is a less reliable inferential strategy than Mill's method of difference. With the method of difference, there is only one culprit to pin the relevant explanation on, for there is by definition only one difference between the two cases. We are thus reasonably sure that this single differentiating feature is causally efficacious. With models as foils, though, because the two cases being compared are the real world and a model world, there will be more than one culprit capable of explaining the relevant difference, for there will be multiple differentiating features present. Here, the theorist needs to give an argument concerning why feature f_n (a member of set *F**) rather than f_m (also a member of set *F**) best explains those properties present in *O* but absent in *O*`. Of course, the theorist's argument cannot show this definitively. Even though she argues there is compelling reason to think it is f_n that

causes the difference, in reality it could be f_m that is doing all the work. Because of this, a theorist who uses models as foils as an inferential strategy is less justified in her conclusion than someone who uses Mill's method of difference. Being less justified does not imply that one lacks any sort of justification at all, though. We do think that models as foils as a method of inference can establish credible beliefs about the world, as is the case in Coase's work.

The whole purpose of this paper is to give an account of the way in which Coase reasons with models. We believe our account of models as foils does this, and we now wish to highlight this. Consider the following reasoning:

1. In Case 1 (the real world), different allocations of property rights result in externalities being resolved in diverging ways (some or all of which may be inefficient).
2. In Case 2 (the model world), all externalities are bargained to the same efficient point regardless of the initial allocation of property rights.
3. In Case 1, we see different allocations of property rights affect the resolution of externalities. This is not so in Case 2.
4. Compared to all the elements present in the real world but absent in the model world, transaction costs are the most plausible explanation as to why allocations of property rights affect the resolution of externalities.
5. As such, there is reason to think that transaction costs influence how different allocations of property rights resolve externalities.

The conclusion stated in (5) is indeed the conclusion that Coase reaches with his model. And, moreover, we believe the structure of reasoning in the schematization is similar to Coase's, at least as we articulated it in §2 above. To put it more informally and less cumbersomely: because property rights wouldn't matter in a world without transaction costs – precisely because these costs are absent – but because we live in a world where both transaction costs are present and property rights *do* matter, we have some reason to believe that transaction costs are the reason *why* these property rights do in fact matter. We have reason to believe this through using the model world as a foil against our real one.

4. Coase and the Austrians

The last section articulated the inferential strategy models as foils, which captures the method of inference Coase uses when he examines a model world without transaction costs. By looking closely at a model world where transaction costs are absent, Coase draws conclusions

about the role transaction costs play in the actual world. We do not think Coase is alone in using this kind of inferential strategy. In particular, we believe several thinkers working within the broadly Austrian school of economics use this same kind of reasoning. In this section, we argue that many working in the Austrian economics tradition interpret competitive equilibrium models using the models as foils approach, focusing on the work of Israel Kirzner first and then F.A. Hayek.[10]

Competitive equilibrium models are models where prices clear markets.[11] While there are many features constituting these models, we highlight only one for our purposes: everyone in the model is a price-taker. There are price-taking producers who try to maximize profits, and price-taking consumers who try to maximize their utility. As Kenneth Arrow and Gerard Debreu characterize it in their classic paper: "It was assumed that each consumer acts so as to maximize his utility, each producer acts so as to maximize his profit, and perfect competition prevails, in the sense that each producer and consumer regards the prices paid and received as independent of his own choices" (Arrow and Debreu 1954: 265).

Just as Coase draws inferences about the real world by examining an unrealistic model world, Kirzner draws inferences about the real world by looking to competitive equilibrium models.[12] That is, Kirzner uses competitive equilibrium models as a foil. Now as we saw, Coase constructs a fictitious world with (*i*) a deeply unrealistic assumption (no transaction costs) and (*ii*) an outcome that deviates from what we observe in our actual world (all externalities are bargained to the same efficient point, regardless of initial allocations of property rights). From the model world with transaction costs conspicuously absent Coase infers the relationship between transaction costs and property law in the actual world. Kirzner, we believe, does something similar.

First, Kirzner notes that the outcomes of competitive equilibrium models are radically different from what we observe in the real world. Namely, stylized markets as depicted in these

[10] Kirzner and Hayek are not alone here. See also Mises (1949/2007: 245-260) on the evenly rotating economy.
[11] While there are different interpretations of perfect competition, the Arrow-Debreu general equilibrium formulation, which *assumes* every agent is a price-taker, dominates economics. For an alternative interpretation of perfect competition – plus commentary on Coase and the Austrian school – see Makowski and Ostroy (2001).
[12] Most economists do not interpret competitive equilibrium models in the manner Kirzner does. For example, Hands (2016: 34) argues that the epistemic value of Arrow-Debreu framework – particularly the work on stability elaborated in Arrow and Hahn (1971) – rose and then fell with its success as derivational robustness.

models are incapable of self-correcting, even though markets do in fact self-correct in the real world. Writes Kirzner:

> Would-be buyers who have been returning home empty-handed (because they have not been offering sufficiently high prices) have *not* learned that it is necessary to outbid other buyers; would-be sellers who return home with unsold goods or resources (because they have been asking prices that are too high) have *not* learned that they must, if they wish to sell, be satisfied with lower prices. Buyers who have paid high prices do not discover that they could have obtained the same goods at lower prices; sellers who have sold for low prices do not discover that they could have obtained higher prices (Kirzner 1973/2013: 11).

This highly unrealistic outcome that deviates from the real world is akin to the outcome produced by Coase's informal model: property rights don't matter, and the same efficient outcome is always reached regardless the underlying property law.

But for Kirzner, what is the equivalent in competitive equilibrium models to Coase's absent transaction costs? That is, what feature is absent from competitive equilibrium models that plausibly explains why market failures go uncorrected? According to Kirzner the issue here is the absence of *entrepreneurs*, who are *not* passive price-takers. Recall what we said above: all competitive equilibrium models assume that everyone is a price-taker. Rather than being mere price-takers, entrepreneurs can seek new ends to maximize and new means by which to maximize these new or existing ends (Kirzner 1963/2011: 16-18). In other words, they do not take prices as given. By introducing new means and new ends, entrepreneurs are capable of correcting market failures. Writes Kirzner:

> Into this imaginary world of men unable to learn from their market experience let us now introduce a group of outsiders who are themselves neither would-be sellers nor would-be buyers, but who *are* able to perceive opportunities for entrepreneurial profits; that is, they are able to see where a good can be sold at a price higher than that for which it can be bought. This group of entrepreneurs would, in our imaginary world, immediately notice profit opportunities *that exist because of the initial ignorance of the market participants* and that have persisted because of their inability to learn from experience (Kirzner 1973/2013: 11).

According to Kirzner, then, the absence of entrepreneurs in competitive equilibrium models is what causes market failures to go uncorrected, just as the absence of transaction costs in Coase's model world causes all externalities to be bargained to the same efficient point. But, clearly, there are entrepreneurs present in the real world and, furthermore, markets self-correct in the real world. This allows Kirzner to conclude that the presence of entrepreneurs in the real world bears some causal relation as to why markets do in fact self-correct in the real world, just as Coase concludes that the presence of transaction costs in the real world is why property law ultimately matters.

We can see more clearly how Kirzner uses models as foils by articulating his argument in terms of the schema used in the prior section. Such reasoning runs roughly as follows:

1. In Case 1 (the real world), markets fail, but also self-correct.
2. In Case 2 (the model world), markets fail, and never self-correct.
3. In Case 1, we see the self-correction of market failures. This is not so in Case 2.
4. Compared to all the elements present in the real world but absent in the model world, the presence of entrepreneurs (i.e., persons who are not mere price-takers) is the most plausible explanation as to why market failures are eventually corrected.
5. As such, there is reason to think that entrepreneurs cause real-world markets to self-correct.

This, of course, is the same style of reasoning used by Coase to draw inductive inferences about the real world from his unrealistic model world. Kirzner does the same thing. He draws inductive inferences about the relationship between entrepreneurs and market failures in the real world by looking to model worlds in which entrepreneurs are conspicuously absent, and the bizarre results that follow from this.

Kirzner is not the only Austrian who uses models as foils. Indeed, Hayek even uses the terminology of *foils*: "For it is only by contrast with this imaginary state, which serves as a kind of *foil*, that we are able to predict what will happen if entrepreneurs attempt to carry out any given set of plans" (Hayek 1941/2009: 32) (emphasis added). Hayek also interprets competitive equilibrium models as foils, but to a different end than Kirzner. In his important paper "The Meaning of Competition," Hayek begins by examining the so-called "perfect competition" assumption that theorists at the time (and still today) typically relied on. According to Hayek's reading, assuming perfect competition is akin to presupposing "complete knowledge of the relevant factors on the part of all participants in the market" (Hayek 1948/2014: 107). When we assume perfect competition, we assume that producers "know the lowest cost at which the

commodity can be produced," and also know "the wishes and desires of the consumers, including the kinds of goods and services which they demand and the prices they are willing to pay" (Hayek 1948/2014: 108). Moreover, "the same situation exists on the side of the consumers or buyers" (Hayek 1948/2014: 108).

Hayek begins by noting something peculiar about this world of perfect information and perfect competition. In particular, though there is so-called perfect competition in the model, none of the activities we typically associate with competition among firms would be present. He writes:

> Now, how many of the devices adopted in ordinary life to that end would still be open to a seller in a market in which so-called 'perfect competition' prevails? I believe that the answer is exactly none. Advertising, undercutting, and improving ('differentiating') the goods or services produced are all excluded by definition—'perfect' competition means indeed the absence of all competitive activities (Hayek 1948/2014: 108-109).

Thus, in the model world of perfect competition, we do not see firms engage in the sorts of *competitive practices* that we typically see in the real world. If everyone has perfect information, there is no reason to advertise or differentiate one's self. The consumer, already knowing the facts, would find such information useless in making her choice.[13]

This absence of competitive practices in perfect competition models can teach us something important about the role such practices play in the real world. Returning to Hayek:

> In actual life the fact that our inadequate knowledge of the available commodities or services is made up for by our experience with the persons or firms supplying them—that competition is in a large measure the competition for reputation or good will—is one of the most important facts which enables us to solve our daily problems. The function of competition is here precisely to teach us *who* will serve us well: which grocer or travel agency, which department store or hotel, which doctor or solicitor, we can expect to provide the most satisfactory solution for whatever particular personal problem we may have to face (Hayek 1948/2014: 109).[14]

[13] See also Hayek (1961/2014: 420) on this point.
[14] For similar comments on the relationship between competition and knowledge, see Hayek (1961/2014: 425).

The basic idea is this: competitive practices are present in real markets *precisely because* they help alleviate the deep knowledge problems that real consumers face. When we have incomplete information about the goods two different firms produce, competitive practices among these firms can help us learn more. This can best be seen by closely examining perfect competition models, where there is a complete dearth of competitive practices. In particular, by seeing what causes the lack of competitive practices in the model (perfect information), we can better understand the role competitive practices play in the real world (ameliorating our ignorance).

Like before, we can summarize Hayek's reasoning with our schema. It goes as follows:

1. In Case 1 (the real world), firms engage in competitive practices.
2. In Case 2 (the model world), firms do not engage in competitive practices.
3. In Case 1 we see firms engage in competitive practices. This is not so in Case 2.
4. Compared to all the elements present in the real world but absent in the model world, consumer ignorance is the most plausible explanation as to why real-world firms engage in competitive practices.
5. As such, there is reason to think that consumer ignorance is why real-world firms engage in competitive practices.

This, of course, is the same style of reasoning that Coase as well as Kirzner uses. Coase draws conclusions about what transaction costs do in the real world by looking at a model world in which transaction costs are absent. Kirzner draws conclusions about what entrepreneurs do in the real world by looking at a model world in which entrepreneurs are absent. And Hayek draws conclusions about what competitive practices do in the real world by looking at a model world in which competitive practices are absent. All three economists use models as foils.

5. The Continuing Relevance of Models as Foils

While we appreciate the importance of economic methodology and the history of economic thought as much as anyone, we argue in this section that understanding Coase's method is a useful tool for practicing economists in everyday research. First, we show that Coase's method of inference has seen a revival within economics without it being recognized. However, it has not been within theory but within the now-dominant paradigm of empirical work: causal inference. Second, we show how Coasean inference can still be used within economic theory to generate new hypotheses.

The dominant paradigm within modern applied economics is causal inference. Instead of studying correlations, the purpose of causal inference is to study the *cause* of some treatment *X* on some outcome variables *Y*. In particular, causal inference within economics follows the "potential outcomes" methodology of Rubin (1974). As Imbens and Wooldridge summarize:

> Rubin proposed the interpretation of causal statements as comparisons of so-called potential outcomes: pairs of outcomes defined for the same unit given different levels of exposure to the treatment, with the researcher only observing the potential outcome corresponding to the level of the treatment received. *Models are developed for the pair of potential outcomes rather than solely for the observed outcome* (Imbens and Wooldridge 2009: 6-7) (emphasis added).

The potential outcomes approach explicitly requires the researcher to specify what could have happened if exposure to the treatment were different. In an experimental setting, the control group can give the researcher the other potential outcome. In non-experimental settings, studying causation is more difficult. It is the job of the researcher to use data and theory to fill in the unobserved potential outcomes. For example, suppose an economist was interested in the causal effect of transaction costs on bargaining outcomes. To do this, she needs to look at potential outcomes. One cannot simply go out into the world and eliminate transaction costs. She could go to the lab (but how do you *really* get rid of transaction costs in the lab?). Or, she could conduct a mental experiment where she imagines a world without transaction costs.

Note that Coase is doing something just like this from a theoretical, rather than empirical, perspective. He wants to know the effects of transaction costs on how property rights resolve externalities. To fully understand the effects of the treatment (transaction costs), Coase must observe potential outcomes. He cannot do this by removing transaction costs from the world – that is impossible – so he builds a mental model of a world where the treatment effect is completely absent. This model world is a potential outcome. Comparing the effects of transaction costs across the space of potential outcomes (the real world and the model world) allows him to draw a causal inference.

While the potential outcomes approach is now standard within economics, it is important to recognize that the exercise is strange.[15] As Cunningham explains:

> To ask questions like this is to engage in a kind of storytelling… What if Bruce Wayne's parents had never been murdered? What if that waitress had won the lottery? What if your friend from high school had never taken that first drink? What if Neo had taken the blue pill? These are the sort of questions that can keep a person up at night. But it's important to note that these kinds of questions are by definition *unanswerable* (Cunningham 2018: 85-6) (emphasis in original).

Coase engages in storytelling when he constructs the unobserved treatment of zero transaction costs and the unobserved potential outcome of efficient bargaining. By comparing this hypothetical treatment to the realized treatment, Coase is arguing for the existence of a causal effect of transaction costs, instead of the more common estimation of the size of an effect.

We have shown how many empiricists are using models as foils without even realizing it; what about contemporary economic theory? Contemporary theorists do not typically use models as foils when they generate hypotheses, but we think this should change. To end, we would like to offer one example detailing how theorists can do this.

Consider the question: what causes coordination failures? A coordination failure is any equilibrium that is Pareto inefficient; if players could coordinate on a different equilibrium, they would all be better off. The usual story is that Althea does not produce hardware because it is worthless without Bertha's software. Bertha does not produce software because it is worthless without Althea's hardware. The outcome without production is self-sustaining, *i.e.* an equilibrium, even if Althea and Bertha prefer the outcome where production occurs. The theorist may come up with several answers for why these coordination failures occur. The typical approach would be for her to write down a model where coordination failure happens, and then explain the aspects of the model that cause the failure. This is not the Coasean approach. The Coasean approach to coordination failures is carried out in Albrecht (2020), who develops a model and uses it as a foil.

[15] Within the field of history, E.H. Carr (1987: 97) famously criticized counterfactuals as a "parlour game with the might-have-beens of history." See Geloso and Glock (forthcoming) for more.

The main theorem in that paper is that when the final goods market is competitive and the producers are price-takers, coordination failure equilibria are not stable and therefore unlikely to exist or persist. The idea is that in a world with price-taking, Althea does not need to worry about what Bertha produces. Instead, Althea only responds to the market price of hardware. If some software producers make small mistakes (because of a trembling hand), then Althea's hardware becomes valuable, and she is led by an invisible hand to produce. This mechanism unravels the coordination failure. For the purpose of our discussion, the key takeaway is that the main theorem proves that coordination failures do *not* occur. However, Albrecht uses this result, from a model where the price system works perfectly, in order to think about the causes of coordination failures in our actual world. He concludes the paper: "If we observe an outcome that looks like a coordination failure, we have a reason to look for the imperfections of competition in the market. When looking for how to use policy, solving imperfections of competition should solve the coordination problem" (Albrecht 2020: 21).

Let us specify more clearly Albrecht's reasoning. In particular:

1. In Case 1 (the real world), coordination failures seem to occur.
2. In Case 2 (the model world), coordination failures would not occur.
3. In Case 1, we see coordination failures. This is not so in Case 2.
4. Compared to all the elements present in the real world but absent in the model world, the presence of imperfections of competition is the most plausible explanation as to why coordination failures occur.
5. As such, there is reason to think that imperfections of competition cause real-world markets to remain in coordination failures.

In other words, Albrecht uses a model of when coordination failures do *not* occur, to draw an inference about their causes when they *do* occur. His conclusion – from a model explicitly designed to eliminate all imperfections of competition – is about situations where there are imperfections. More serious empirical work will be required to measure how important the imperfections of competition are to coordination failures in the real world. But what Albrecht does, through a Coasean inference, is generate new hypotheses about causal mechanisms in the world. He does this by using his model as a foil.

Works Cited

Albrecht, Brian C. 2020. "Investment without Coordination Failures." Working Paper.

Allen, Douglas W. 2015. “The Coase Theorem: Coherent, Logical, and Not Disproved.” *Journal of Institutional Economics* 11: 379-390.

Arrow, Kenneth J. and Gérard Debreu. 1954. “Existence of an Equilibrium for a Competitive Economy.” *Econometrica* 22: 265-290.

Arrow, Kenneth. J. and Frank H. Hahn. 1971. *General Competitive Analysis*. Amsterdam: Elsevier.

Aslanbeigui, Nahib and Steven G. Medema. 1998. “Beyond the Dark Clouds: Pigou and Coase on Social Cost.” *History of Political Economy* 30:601-625.

Bertrand, Elodie. 2010. “The Three Roles of the ‘Coase Theorem’ in Coase’s Works.” *European Journal of the History of Economic Thought* 17: 975-1000.

Bertrand, Elodie. 2015. “‘The Fugitive’: The Figure of the Judge in Coase’s Economics.” *Journal of Institutional Economics* 11: 413-435.

Bertrand, Elodie. 2016. “Coase’s Choice of Methodology.” *Cambridge Journal of Economics* 40: 1203-1227.

Boettke, Peter J. 1997. “Where Did Economics Go Wrong? Modern Economics as a Flight from Reality.” *Critical Review* 11: 11-64.

Boettke, Peter J. 2012. *Living Economics: Yesterday, Today, and Tomorrow*. Oakland: Independent Institute.

Carr, Edward H. 1987. *What is History?* Middlesex: Penguin UK.

Coase, Ronald H. 1960/1988. “The Problem of Social Cost,” in *The Firm, the Market, and the Law*: 95-156. Chicago: University of Chicago Press.

Coase, Ronald H. 1988a. “Notes on the Problem of Social Cost,” in *The Firm, the Market, and the Law*: 157-185. Chicago: University of Chicago Press.

Coase, Ronald H. 1988b. “The Firm, the Market, and the Law,” in *The Firm, the Market, and the Law*: 1-32. Chicago: University of Chicago Press.

Cunningham, Scott. 2020. *Causal Inference: The Mixtape, V. 1.8.* Available at: https://scunning.com/cunningham_mixtape.pdf.

Frischmann, Brett M. and Alain Marciano. 2015. “Understanding *The Problem of Social Cost*.” *Journal of Institutional Economics* 11: 329-352.

Geloso, Vincent and Judge Glock. Forthcoming. “The New History of Capitalism and the Methodologies of Economic History.” *Essays in Economic and Business History*.

Hands, D. Wade. 2015. “Derivational Robustness, Credible Substitute Systems and Mathematical Economic Models.” *European Journal of Philosophy of Science* 6: 31-53.

Hayek, F.A. 1941/2009. *The Pure Theory of Capital*. Auburn, Alabama. The Ludwig von Mises Institute.

Hayek, F.A. 1948/2014. “The Meaning of Competition.” In *The Collected Works of F.A. Hayek Volume Fifteen*, edited by Bruce Caldwell: 105-116. Chicago: University of Chicago Press.

Hayek, F.A. 1961/2014. “A New Look at Economic Theory.” In *The Collected Works of F.A. Hayek Volume Fifteen*, edited by Bruce Caldwell: 374-426. Chicago: University of Chicago Press.

Hsiung, Bingyuan. 2004. “An Interpretation of Ronald Coase’s Analytical Approach.” *History of Economics Review* 39: 12-32.

Imbens, Guido W. and Jeffrey W. Wooldridge. 2009. “Recent developments in the econometrics of program evaluation." *Journal of Economic Literature* 47:5-86.

Kirzner, Israel M. 1963/2011. *Market Theory and the Price System*. Indianapolis: Liberty Fund.

Kirzner, Israel M. 1973/2013. *Competition and Entrepreneurship*. Indianapolis: Liberty Fund.

Makowski, Louis and Joseph M. Ostroy. 2001. “Perfect Competition and the Creativity of the Market.” *Journal of Economic Literature* 39: 479-535.

McCloskey, Deidre. 1998. “The So-Called Coase Theorem.” *Eastern Economic Journal* 24: 367-371.

Medema, Steven G. 1994. “The Myth of Two Coases: What Coase Is Really Saying.” *Journal of Economic Issues* 28: 208-217.

Medema, Steven G. 1995. “Ronald Coase on Economics and Economic Method.” *History of Economics Review* 24: 1-22.

Medema, Steven G. 2011a. “A Case of Mistaken Identity: George Stigler, ‘The Problem of Social Cost,’ and the Coase Theorem.” *European Journal of Law and Economics* 31: 11-38.

Medema, Steven G. 2011b. “HES Presidential Address: The Coase Theorem Lessons for the Study of the History of Economic Thought.” *Journal of the History of Economic Thought* 33: 1-18.

Medema, Steven G. 2015. “‘A Magnificent Business Prospect…’ the Coase Theorem, the Extortion Problem, and the Creation of Coase Theorem Worlds.” *Journal of Institutional Economics* 11: 353-378.

Mill, John Stuart. 1843. *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation.* eBook reproduced by Gale, Cengage Learning.

Mireles-Flores, Luis. 2018. “Recent Trends in Economic Methodology: A Literature Review.” *Research in the History of Economic Thought and Methodology* 36A: 93-126.

Mises, Ludwig von. 1949/2007. *Human Action: a Treatise on Economics*. Indianapolis: Liberty Fund.

Rubin, Donald B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66:688-701.

Stigler, George. 1966. *The Theory of Price*. 3rd Edition. New York: Macmillan.

Wallis, John J. and Douglass North. 1986. “Measuring the Transaction Sector in the American Economy, 1870-1970.” In *Long-Term Factors in American Economic Growth*, edited by Stanley L. Engerman and Robert E. Gallman: 95-162. Chicago: University of Chicago Press.