A Note on the Semantics of Linear Regression

Raja Panjwani*

December 23, 2024

Abstract

The use of linear regression is ubiquitous across the social and behavioral sciences, and yet researchers rarely hold that the variables in their target systems are in fact linearly related. This raises the question of how to interpret linear regression coefficients when there is 'functional misspecification' and the target system exhibits nonlinearity. Here, recent methodological discussions among practitioners have mirrored philosophical debates over scientific realism. In this paper, I frame the prevailing views in terms of their stance on the scientific realism debate. I then present a novel realist interpretation of linear regression - the 'secant interpretation' - based on a property I derive about the manner in which linear regression coefficients represent properties of non-linear systems of interest.

1 Introduction

Suppose a researcher is interested in a target system consisting of two continuous random variables X and Y.[1] They take for granted that there exists some objective probability distribution f_X over the values of X, f_Y over the values of Y, and a joint distribution $f_{X,Y}$ over their joint values. However, these distributions are unknown to the researcher, so they are unable to calculate their quantity of interest, which is the conditional expectation function, $\mathbb{E}[Y|X]$. Following meticulous data gathering protocols, the researcher proceeds

*Department of Biology, Georgetown University, Washington, DC, 20057, U.S.A., rp1223@georgetown.edu

[1] I follow much of the relevant literature which confines attention to the case of univariate models; however, I will discuss how the results of the paper can be generalized to the multivariate case. The regression which is perhaps most frequently performed in practice is the calculation of a stock's β, which is the slope coefficient from a linear regression of a stock's returns on the corresponding market returns. Several regressions such as this one are primarily concerned with correlation rather than causation; I only briefly discuss causality in Section 5.

to collect from the population an i.i.d sample S=$\langle(x_1,y_1),...,(x_N,y_N)\rangle$, then performs a linear regression which identifies ($\hat{\beta}_0$,$\hat{\beta}_1$) such that:

$$(\hat{\beta}_0, \hat{\beta}_1) = \underset{(b_0,b_1)\in\mathbb{R}^2}{\operatorname{argmin}} \mathbb{E}[(y_i - b_0 - b_1 x_i)^2] \tag{1}$$

Here, the expectation is taken over the empirical distribution of X. As long as the researcher resamples (x_i,y_i) pairs from the population in each experiment,[2] the estimators ($\hat{\beta}_0$, $\hat{\beta}_1$) are unbiased estimates of the population linear regression parameters (β_0,β_1):[3]

$$(\beta_0, \beta_1) = \underset{(b_0,b_1)\in\mathbb{R}^2}{\operatorname{argmin}} \mathbb{E}[(Y - b_0 - b_1 X)^2] \tag{2}$$

This time, the expectation is taken over the true (unknown) probability distribution, f_X. Thus, the researcher who was originally interested in the functional relationship between X and Y, imposed a linear model on a dataset of their values,[4] and is now in possession of a slope coefficient $\hat{\beta}_1$. They realize that this coefficient is an unbiased estimate of β_1, but they wonder what β_1 represents about the true relationship between X and Y, because they have no *a priori* reason to believe that the true relationship is even approximately linear.[5]

In this article, I outline the prevailing interpretive accounts, and I classify them by the extent to which they offer a realist interpretation of β_1. I then present a novel realist interpretation which I call the 'secant interpretation', based on a property I derive about the manner in which linear regression lines intersect with non-linear conditional expectation functions. The two main upshots are first, a concise realist semantics for misspecified linear regression coefficients, and second, an epistemically intriguing 'possibility result': the theorem implies that although the true functional relationship between the variables of interest is unknown, the methodology of linear regression (in particular, the fact that regression minimizes loss) guarantees (under very weak conditions) that β_1 represents a rate of change along the true conditional expectation function.

[2]Of course, this is often impossible in practice; for example, the target system may be non-stationary.

[3]However, Buja et al (2015) show that the typical textbook treatment of linear regression with fixed-covariates, i.e., resampling only y_i for a fixed set of covariate values, will be biased when the true response function is nonlinear.

[4]The choice of linear model can be guided by so-called 'regression diagnostics' such as misspecification tests. However, Berk et al (2018) argue convincingly that these diagnostics underdetermine whether it is the linearity assumption, or other assumptions - for example distributional assumptions - which point to misspecification. In practice, linear regression is relied upon without extensive use of misspecification tests.

[5]For example, "the use of linear regression in political science ... rarely follows from theories of social and political processes implying that a linear functional form between y and X holds globally" (Beck and Jackman 1995, p.598).

2 Prevailing Interpretations

Scientific realists are "committed to the mind-independent existence of the world investigated by the sciences...[and] a literal interpretation of scientific claims about the world."[6] Much of the statistical methodology literature already denies the first of these commitments. Leamer (Hendry et al 1990) and Keuzenkamp (2000) deny that an objective probability distribution generating the data (called the 'data generating process', or DGP) exists in the first place: "...The DGP does not exist. It is an invention of our mind."[7] However, even taking for granted that there is a true DGP, there is still an interpretive challenge for the realist based on the second commitment, as the semantics of the linear model appear to be literally false when the true conditional expectation function is non-linear.

> The issue is that, in the presence of nonlinearity, slopes lose their usual interpretation: [β_1] is no longer the average difference in Y associated with a unit difference in X[.] The challenge is to provide an alternative interpretation that remains valid and intuitive.
>
> \- Buja et al (2015), p.20.

Simply put, even if the realist is given the true value and units of β_1, it is still unclear what they have learned about the underlying target system consisting of X and Y, in light of the fact that the true conditional expectation function is assumed to be non-linear.

The instrumentalist has a much easier interpretive task, as they are unencumbered by the realist commitment to interpret their models as mind-independent representations of their target system. There is a straightforward instrumentalist interpretation of linear regression in terms of betting behavior: although it may be literally false that $\mathbb{E}[Y|X]=\beta_0+\beta_1$X, if one were forced to bet on $\mathbb{E}[Y|X]$ and constrained to use only linear models then the linear model with coefficients (β_0,β_1) is the 'best linear predictor' (BLP), where 'best' is understood in terms of minimizing expected squared loss. An important attempt to ground scientific realism in terms of betting behavior is Dennett's 'real patterns' argument (Dennett 1991). Dennett considers a pattern to be 'real' if predictions on the basis of that pattern perform better than pure randomness. However, Ross et al (2009) argue convincingly that Dennett's criterion is a necessary, but not sufficient condition for realism. Interpreting coefficients in terms of their use in betting strategies offers a *pragmatics* of linear regression, but the realist seeks a *semantics* of the model which interprets its expressions as designating intrinsic properties of the target system, accounting for *why* the model is useful in certain contexts.

[6] Chakravartty (2017).

[7] Keuzenkamp 2000, p.56. See also Edward Leamer: "I ... don't think there is a true data generating process..." (Hendry et al 1990, p.188).

Several practitioners have settled on a perspectivalist interpretation of regression (e.g., Hoover 2011, 2019).[8] In his initial formulation of perspectivalism, Giere (Giere 2006) points to the Mercator projection map as an example of a perspectival model: one can learn *about* the globe by studying an atlas of its projections, even though each projection gives rise to different distortions. There is an analogous 'projection' view in the statistical methodology literature:[9]

> If one thinks of the parameter of interest as the true conditional mean of Y, given the covariates of interest, then one is almost always wrong. However, if one views the parameter of interest as a projection of the "truth" onto some smaller set of models (eg, linear), this is often a reasonable parameter of interest. ...the projection of the population average is something one can hope to estimate, and ... it bears a rigorously definable relation to the true underlying association.
>
> - Hubbard et al 2010, p.471-472

There is, however, an important disanalogy between these cases. In the Mercator projection map case, we have access to the target system (the Earth), so it is possible to identify properties which the map 'gets right' about its target, for example various topographical properties. In the case of linear regression, by contrast, the target system is assumed to be unknown, which makes it difficult to take a realist stance towards a parameter like β_1 in terms of what it even approximately designates, or corresponds to in the target system.

> ...it is usually impossible to know whether the regression model specified by the analyst is the means by which the data were generated. A common fallback, therefore, is to claim that the model specified is "close enough." But there is no way to know what "close enough" means. One requires the truth to quantify a model's disparities from the truth, and were the truth known, there would be no need to analyze any data.
>
> - Berk et al 2018, p.637

Some perspectivalists (e.g., Hoover 2011) have denied that there is a *perspectiveless* specification of the target system in the first place, and holding models to the standard of an

[8] So as to not take a stand on whether perspectival realism is a genuine form of realism (Teller (2019) argues that it is, whereas Panagiotatou and Psillos (2023) argue that it is not) I use the neutral term 'perspectivalism'.

[9] This perspective has recently been advocated in a series of papers by statisticians at the Wharton School (Berk et al 2014, 2018, 2021, Buja et al 2015): "The researcher forgoes trying to estimate the true conditional means and settles for trying to estimate the best linear approximation of that truth." (Berk et al 2018, p.643).

idealized ‘perfect model’ (Teller 2001) is too high of a bar for realism to meet. Deciding the issue of whether there is a ‘perfect model’ is outside the scope of this paper: in the context of our regression model the question boils down to whether there is a true, objective joint probability distribution over the random variables X and Y, or whether the probability distribution is ‘subjective all the way down’. In this paper we will follow much of the literature in taking for granted that there is a true DGP, and our interpretive challenge will be to identify a realist semantics for what the linear regression parameter β_1 refers to in the true conditional expectation function.

Turning now to realist interpretations, once one rejects a naïve realism which assumes that the DGP does in fact have linear conditional expectations, there has been only one straightforwardly realist interpretation in the literature. This is the so-called ‘weighted average derivative’ interpretation (Angrist and Krueger 1999) which is based on the following theorem.[10]

Theorem 1. Let X be a continuous random variable with density f_X which has full support over a closed interval [a,b], and let $\mathbb{E}[Y|X]$ be continuously differentiable over the support of X. Then the slope parameter of the BLP can be written as:

$$\beta_1 = \mathbb{E}\left(w(X)\frac{d}{dx}\mathbb{E}[Y|X]\right) \tag{3}$$

$$w(X) = \frac{1}{f_X(x)}\frac{\mathbb{E}[X-\mu_X|X \geq x](1-F_X(x))}{\int_{v=a}^{v=b}\mathbb{E}[X-\mu_X|X \geq v](1-F_X(v))\,dv} \tag{4}$$

Thus, the parameter β_1 can be interpreted as a weighted average of tangent vectors to the true conditional expectation function, where the derivatives with x-values closer to their mean μ_X are given greater weight.[11] There is a straightforward sense in which this interpretation is realist: the misspecified linear regression parameter is a *weighted average of the truth*.

Although the average derivative interpretation is realist, there are reasons to seek a more intuitive account. The weight function is rather complex, and does not extend to higher dimensions in an easily interpretable manner, as noted by Angrist and Krueger: “in multivariate regression models ... this interpretation is complicated by the fact that the OLS slope vector is actually matrix-weighted average of the gradient of the CEF. Matrix-weighted averages are difficult to interpret[.]”[12] Furthermore, realists strive to “take theoretical statements at face value” (Chakravartty 2017), and regression coefficients are in practice interpreted as

[10] See Graham et al 2010, p.37-38 Lemma A1 for a proof.

[11] Gelman and Park (2008) present an analogous finite sample interpretation of $\hat{\beta}_1$ based on a weighted average of slope vectors instead of derivatives.

[12] Angrist and Krueger 1999, p.1311.

estimates of rates of change between random variables. By interpreting a coefficient as an estimate of an average derivative across the entire support of the regressor space, it masks the extent to which the coefficient estimates a particular rate of change, with a restricted domain of validity. In the next section, it is shown that under very weak additional constraints on the above setup (which are very likely to be satisfied by physical systems of interest), it is possible to derive a realist semantics which retains the intuition that a linear regression coefficient estimates a rate of change between the variables in a target system, even in higher dimensions.

3 Broken Clock Theorem

The secant interpretation is grounded in the theorem of this section, where it is shown that, possibly aside from knife-edge situations, a linear regression line will intersect with its (non-linear) conditional expectation function at least twice, justifying its interpretation as a secant through the true conditional expectation function. I call this result the 'broken clock theorem' as a tongue in cheek reference to the fact that a broken clock is also right twice a day.

Theorem 2. Let X be a continuous random variable with density f_X which has full support over a closed interval [a,b], and let $\mathbb{E}[Y|X]$ be twice differentiable over the support of X. Define the loss function L(X)=$\mathbb{E}[Y|X]$-BLP(X). If there is no unique x*∈[a,b] such that L(X=x*)=L′(X=x*)=L″(X=x*)=0, then $\exists x_1,x_2 \in$[a,b] $(x_1 \neq x_2)$: L(x_1)=L(x_2)=0.

Proof. We will make use of the following property of best linear predictors.

Lemma 1. $\mathbb{E}$[L(X)]=0.

Proof. Recall that the first-order conditions of linear regression imply that:

$$\beta_0 = \mathbb{E}[Y] - \mathbb{E}[X]\frac{Cov(X,Y)}{Var(X)} \tag{5}$$

$$\beta_1 = \frac{Cov(X,Y)}{Var(X)} \tag{6}$$

Taking the expected value of L(X) with respect to X and applying the law of iterated expectations, we get:

$$\mathbb{E}_X[L(X)] = \mathbb{E}_X\mathbb{E}_Y[Y|X] - \mathbb{E}[Y] + \frac{Cov(X,Y)}{Var(X)}\mathbb{E}_X[\mathbb{E}_X[X] - X] \tag{7}$$

$$= \mathbb{E}[Y] - \mathbb{E}[Y] + \frac{Cov(X,Y)}{Var(X)}[\mathbb{E}_X[X] - \mathbb{E}_X[X]] = 0 \tag{8}$$

□

We now proceed by contradiction, i.e., by showing that there cannot be exactly zero or one intersection between $\mathbb{E}[Y|X]$ and BLP(X).

Case 1: Suppose there are zero intersections. Then it is either the case that for all x∈[a,b], $\mathbb{E}[Y|X=x]$>BLP(X=x) or it is the case that for all x∈[a,b], $\mathbb{E}[Y|X=x]$<BLP(X=x). In either case, the expected loss will be nonzero, contradicting the lemma.

Case 2a: Suppose there is exactly one intersection, i.e., a unique x*∈[a,b] for which L(X=x*)=0. Then, L′(X=x*) is either zero or nonzero. If L′(X=x*)=0, then by assumption L″(X=x*)≠0, which means BLP(X) is tangent to $\mathbb{E}[Y|X]$ at X=x*. But then since f_X is non-degenerate in [a,b], $\mathbb{E}$[L(X)]≠0, contradicting the lemma (note that x*≠a and x*≠b for the same reason).

Case 2b: Suppose there is exactly one intersection at X=x* with L′(X=x*)≠0. Without loss of generality, suppose L′(X=x*)>0, so that BLP(X)>$\mathbb{E}[Y|X]$ for x<x* and BLP(X)<$\mathbb{E}[Y|X]$ for x>x*. Then since L(X) is twice differentiable, it is continuously differentiable, so there exists an ϵ>0 such that within the open ball B_ϵ(x*), L(X) is strictly increasing (see Tao 2015, p.152). Now divide [a,b] into [a,x*-ϵ] (region 1), B_ϵ(x*) (region 2), and [x*+ϵ, b] (region 3). We will construct a rotated line, through the point (x*, BLP(X=x*)), which has lower expected squared loss than BLP(X), contradicting the assumption that BLP(X) minimizes expected squared-loss.

Since L(X) is strictly increasing in B_ϵ(x*), there exists a counterclockwise rotation of BLP(X) around X=x* by angle θ, called R_θ, such that, for x>x* in region 2, BLP(X)<R_θ<$\mathbb{E}$[Y|X] and for x<x* in region 2, $\mathbb{E}$[Y|X]<R_θ<BLP(X). Subtracting BLP(X) from the right hand side of both inequalities and using the fact that the quadratic function preserves the order of positive real numbers, it follows that R_θ reduces squared loss for all x in region 2, and thus reduces expected squared loss in region 2. Furthermore, any line in the sequence converging to BLP(X), $\{l_n\}$ defined by $l_1 = R_\theta$ and l_n=$R_{\frac{\theta}{n}}$ for n∈ℕ will reduce expected squared loss in region 2. Thus, if we can show that an element of $\{l_n\}$ also reduces expected squared loss in regions 1 and 3, we will be done.

Given that |L(X)| is continuous over compact intervals in regions 1 and 3, it achieves a min value over these regions, and suppose that min value has magnitude δ>0. Since $\{l_n\}$ converges pointwise to BLP(X) as n→∞, there exists some M∈ℕ such that for all N>M, and all x in regions 1 and 3, $|l_N(X=x)$-BLP(X=x)|<δ. We then have $\mathbb{E}$[Y|X]<$l_N(X)$<BLP(X) for x<x* (in both regions 1 and 2) and BLP(X)<$l_N(X)$<$\mathbb{E}$[Y|X] for x>x* (in both regions 2 and 3). Thus, any such $l_N(X)$ has lower expected squared loss than BLP(X), contradicting that BLP(X) is expected squared loss minimizing, which completes the proof. $\square$

The condition that there is no unique x*∈[a,b] such that L(X=x*)=L′(X=x*)=L″(X=x*)=0 implies that there is no unique intersection with a saddle point of the loss function at that value. Violations of this condition would be knife-edge situations which are unlikely to be seen in practice. Note that the theorem provides a sufficient condition for two intersections; even if there is a saddle-point where loss is zero, there may still be at least two intersections. For our purposes, however, we consider the condition sufficiently knife-edge that we can in practice take it for granted that the BLP is a secant through the true conditional expectation function.

Corollary 1. $\exists x_l$∈(a,b): $\frac{d}{dx}\mathbb{E}[Y|X = x_l]$=$\beta_1$.

This corollary, which is an elementary consequence of the mean value theorem, implies that the slope of the BLP represents an instantaneous rate of change *somewhere* along the true conditional expectation function.

4 Secant Interpretation

The broken clock property of linear regression allows for an intuitive semantics for the slope parameter in the univariate case:

$$\exists x \exists x' : \beta_1 = \frac{\mathbb{E}[Y|X = x'] - \mathbb{E}[Y|X = x]}{x' - x} \tag{9}$$

This interpretation extends to the multivariate case in a straightforward way. Since a linear regression (hyper)plane will always be a *secant plane* through the true conditional expectation surface[13] a parameter in a linear regression model can be interpreted as the contribution of the corresponding regressor to the expected change in the outcome variable between two points in the regressor space where these functions intersect. More precisely, if the multivariate linear regression model is given by $\mathbb{E}[Y|X] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_N X_N$ but the true conditional response surface is assumed to be non-linear, then the broken clock property entails that:

$$\exists x \exists x' : \mathbb{E}[Y|X = x'] - \mathbb{E}[Y|X = x] = \beta_1(x'_1 - x_1) + ... + \beta_N(x'_N - x_N) \tag{10}$$

[13] A secant plane to a non-linear surface intersects with a non-trivial section of the surface, as opposed to a tangent plane which may only 'slice' a non-linear surface at a point. A linear regression hyperplane cannot be a tangent plane to a non-linear conditional expectation surface because then the expected loss will be non-zero, violating the lemma of the previous section which holds in higher dimensions (see Theorem 4.1.5 in Aronow and Miller 2019, p.149).

Thus, in the multivariate case, a given parameter β_1 is also interpreted in terms of an existence claim: there exist two points in the regressor space $(x_1,...,x_N)$,$(x'_1,...,x'_N)$ such that the shift from x_1 to x'_1 contributes $\beta_1(x'_1$-$x_1)$ to the expected change in the outcome variable, additively with the corresponding changes in the other regressors. Crucially, one cannot meaningfully compare the magnitudes of the regression parameters without knowing the magnitudes of the corresponding shifts in the regressor values. Since the true conditional expectation function is unknown, so are these regressor values.

Notwithstanding this, the main upshot of the broken clock theorem is philosophical: the researcher can have confidence that their slope coefficient is an unbiased estimate of a parameter which 'latches onto' a true rate of change along the target system, even though that target system is assumed to be unknown. It is often said that linear regression models are 'linear approximations'; however, not all linear approximations have the broken clock property. To take an example from physics, pre-relativistic momentum p=mv is a linear approximation to relativistic momentum p=$\frac{mv}{\sqrt{1-\frac{v^2}{c^2}}}$, when v is small compared to c. However, these functions only intersect once, at v=0, so the pre-relativistic momentum function cannot be interpreted as a secant through the relativistic momentum function. Linear regression, by virtue of the fact that it minimizes loss, is a special kind of linear approximation: one which necessarily characterizes a rate of change along its target system.

5 Discussion

Following Suppes (1962), philosophers of science have distinguished between models of (scientific) *theories* and models of *data* (Frigg & Hartmann 2021). The scientific realism literature has primarily been concerned with models of theories: debates persist about, for example, which forms of 'selective realism' about model-elements are resistant to challenges from 'pessimistic metainduction' from theory change in science (e.g., Harker 2013). In contrast, philosophical discussions of data models have thus far focused on methodological issues arising from data manipulation (Harris 2003) and curve-fitting (Forster & Sober 1994), but without a corresponding search for realism about elements of data models. Interestingly, it is practitioners themselves who have advocated for clarity about model-semantics:

> ...it is easy to agree with G.E.P. Box's famous dictum, ["all models are wrong, but some are useful" (Box 1979)] but there are real consequences that cannot be ignored or minimized by hand waving ... [as it] affects how we interpret estimates and how we obtain valid statistical inferences and predictions.

- Berk et al 2021, p.83[14]

As an example of the "real consequences" of these interpretive challenges, it is worth comparing the semantics of the aforementioned average derivative interpretation with the semantics of the secant interpretation. Consider a univariate linear regression model which estimates the *causal* effect of X on outcome variable Y, and suppose the true relationship is non-linear (i.e., there is functional misspecification, but suppose no omitted variable bias). The researcher obtains a single linear regression slope, whose semantics on the average derivative interpretation are as follows:

> ...the slope is an average. Consequently, it will likely overstate or understate the slope at any particular observation. ...Because of these limitations, any step from estimation to causal inference will be challenging. There are the usual conceptual issues such as how to map ... regression coefficients to real-world manipulations of causal variables. But in addition, under what circumstances does it make sense to use linear approximations as the basis for any causal claims? ...the issues are complicated[.]
>
> - Berk et al 2014, p.17

Here we can see a drawback of an insufficiently realist semantics: interpreting linear regression coefficients in terms of average derivatives is mathematically correct, but it obscures the ways in which the model represents causal features of the physical target system. On the secant interpretation, however, the semantics are straightforward: the coefficient is an unbiased estimate of the rate of change in the conditional expectation of Y *caused* by a hypothetical intervention between two X-values. Of course, *which* X-values remains unknown (the broken clock theorem only guarantees their existence), but the semantics are clear: the slope corresponds to a real causal relationship in the target system.

This example illustrates that there are interpretive consequences of identifying a realist semantics for data-models. When compared with models of theories, data models present different challenges and opportunities in the search for selective realism about model-elements. Models of theories typically explain and predict empirical data using theoretical constructs (e.g., utility functions, fitness functions, wave-functions), and the challenge for the realist is

[14] The Box dictum has always been in tension with the no-miracles argument for scientific realism (Putnam 1975), whereby the fact that a model is useful is an indication that it is 'getting something right' about its target system. In contrast, this paper has proceeded in the reverse direction: the broken clock theorem implies that 'all linear models get something right', and yet it is premature to go as far as Angrist and Pischke (2009, p.xii) in stating that "linear regression provides useful information about the conditional mean function regardless of the shape of this function." Rather, we are at most warranted in the conclusion that 'all linear regression models get something right, but only some are useful'.

to justify their belief in the theoretical constructs in light of the fact that theory change is inevitable whereas the preservation of theoretical constructs in successor theories is not. On the other hand, data models are generated algorithmically from the data itself (e.g., loss-minimization in the case of linear regression), and their targets are properties of the true joint distribution of the variables in the data (e.g., rates of change in the true conditional expectation function). In data models, theoretical constructs (e.g., regression coefficients) are couched in the vocabulary of the target system itself, which offers new prospects for the realist seeking to determine how those theoretical constructs correspond to properties in the target system.

In this paper, the functional relationship between the variables in the target system was assumed to be unknown, and an existence theorem was used to demonstrate that the *method* of linear regression guarantees that one of its theoretical constructs (the β-parameter) represents a rate of change in the target system, regardless of its true functional form. The result can be understood as a proof of concept that since data models are formally generated, and because their theoretical constructs are given in the vocabulary of their target system, realism about elements of data models can sometimes be formally demonstrated. This perspective offers new opportunities for a realist justification of scientific practices. For example, a common practice in regression modeling involves modifying functional forms (e.g., by adding interaction terms, or non-linear terms) to see whether coefficients remain significant. Assuming model misspecification, a realist justification for such practices would seek to identify conditions on the relationship between a model and its target system under which such inferences are sound.

6 Conclusion

The interpretation of linear regression coefficients under functional misspecification presents a challenge for the scientific realist. The challenge is to provide a semantics whereby the linear relationship modeled refers to a property of the underlying target system. This paper identifies such a correspondence via the secant interpretation: the derived 'broken clock property' allows the realist to interpret a linear regression coefficient as estimating a rate of change along the true conditional expectation function. Coefficients therefore retain their usual interpretation as slopes, though with a restricted domain of applicability.

Discussions of scientific realism about model-elements have typically centered around models of scientific *theories* to the exclusion of data models. In the case of linear regression, the model-element of interest is the population regression parameter, and the broken clock theorem demonstrates that this parameter can be identified as an intrinsic property of the target system, even though the specific functional form of the target system is assumed to be unknown. The secant interpretation is thus a form of selective realism, and offers a proof of concept that data model realism can sometimes be formally demonstrated, with interpretive implications for scientific practice.

Acknowledgements I am grateful to PM Aronow, William Greene, and Cyrus Samii for valuable feedback on earlier drafts of this paper.

References

1. Angrist, J.D. and Krueger, A.B. " Empirical Strategies in Labor Economics", Princeton University and NBER (1999).

2. Angrist, J. D. and Pischke, J.S. "Mostly Harmless Econometrics", Princeton: Princeton University Press (2009).

3. Aronow, P.M. and Miller, B.T. "Foundations of Agnostic Statistics", Cambridge: Cambridge University Press (2019).

4. Beck, N., and Jackman, S. "Beyond Linearity by Default: Generalized Additive Models." American Journal of Political Science (1995) 42(2):596-627.

5. Berk, R., Brown, L., Buja, A., George, E., Pitkin, E., Zhang, K., Zhao, L., "Misspecified Mean Function Regression: Making Good Use of Regression Models That Are Wrong", Sociological Methods and Research, (2014) p.1-30.

6. Berk, R., Brown, L., Buja, A., George, E., Zhao, L., "Working with Misspecified Regression Models", Journal of Quantitative Criminology, Vol. 34 (2018) p.633-655.

7. Berk, R., Buja, A., Brown, L., George, E., Kuchibhotla, A.K., Su, W., Zhao, L., "Assumption Lean Regression", The American Statistician, Vol. 75 (1) (2021) p.76-84.

8. Box, G., "Robustness in the strategy of scientific model building", in Launer, R.L.; Wilkinson, G.N., (eds.) Robustness in Statistics, Academic Press (1979) p. 201-236.

9. Buja, A., Berk, R. A., Brown, L. D., George, E. I., Pitkin, E., Traskin, M., Zhao, L., and Zhang, K. "Models as Approximations - A Conspiracy of Random Regressors and Model Deviations Against Classical Inference in Regression". Statistical Science (2015) p.1-44.

10. Buja, A., Brown, L. D., Kuchibhotla, A.K., Berk, R. A., George, and Zhao, L., "Models as Approximations II - A Model-Free Theory of Parametric Regression". Statistical Science Vol.34, No.4 (2019) p.545-565.

11. Chakravartty, A., "Scientific Realism", The Stanford Encyclopedia of Philosophy (Summer 2017 Edition), Edward N. Zalta (ed.)

12. Dennett, D., "Real Patterns", The Journal of Philosophy, (1991). Vol.88(1): p.27-51.

13. Forster, M., and Sober, E., "How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions", The British Journal for the Philosophy of Science, Vol.45, No.1 (1994), p.1-35.

14. Frigg, R., and Hartmann, S., "Models in Science", The Stanford Encyclopedia of Philosophy (2nd Edition), Edward N. Zalta (ed.). Online at http://plato.stanford.edu/entries/models-science/

15. Gelman, A. and Park, D. K., "Splitting a Predictor at the Upper Quarter or Third and the Lower Quarter or Third", The American Statistician (2008). Vol. 62(4).

16. Giere, R.N. "Scientific Perspectivism". Chicago: University of Chicago Press (2006).

17. Graham, B., Imbens, G., Ridder, G., "Measuring the Effects of Segregation in the Presence of Social Spillovers: A Nonparametric Approach", NBER Working Paper (2010).

18. Harker, D., "How to Split a Theory: Defending Selective Realism and Convergence without Proximity", The British Journal for the Philosophy of Science, Vol.64 (2013), p.79-106.

19. Harris, T., "Data Models and the Acquisition and Manipulation of Data", Philosophy of Science, Vol 70, No.5 (2003), p.1508-1517.

20. Hendry, D.F. Leamer, E.F., and Poirier D.J. "The ET Dialogue: A Conversation on Econometric Methodology". Econometric Theory (1990) 6(2): 171-261.

21. Hoover,K. "Pragmatism, perspectival realism, and econometrics", Economic Research Initiatives at Duke, No.88 (2011).

22. Hoover,K. "Models, truth, and analytic inference in economics", CHOPE Working Paper No.2019-01, Duke University (2019).

23. Hubbard, A., Ahern, J., van der Laan, M., Bruckner, T., “To GEE or Not to GEE.” Epidemiology (2010), Vol.21 (4) p.467-474.

24. Keuzenkamp, H.A., “Probability, Econometrics and Truth”. Cambridge University Press (2000).

25. Panagiotatou, M., Psillos, S., “Perspectival Realism”, Notre Dame Philosophical Reviews (2023).

26. Putnam, H., “Mathematics, Matter and Method”, Cambridge: Cambridge University Press (1975).

27. Ross, D., Ladyman, J., Collier, J., “Rainforest Realism and the Unity of Science”, Everything Must Go: Metaphysics Naturalized Chapter 4, Oxford University Press (2009).

28. Suppes, P., “Models of Data”, in Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress. Stanford University Press, p.252-261 (1962).

29. Tao, T., “Analysis II (Third Edition)”, Springer Texts and Readings in Mathematics, Vol.38 (2015).

30. Teller, P., “Twilight of the Perfect Model Model”, Erkenntnis Vol.55 (3) (2001), p.393-415.

31. Teller, P., “What is Perspectivism, and Does It Count as Realism?”, Understanding Perspectivism (2019), p.49-64.