NBER WORKING PAPER SERIES

A SIMPLE RATIONAL EXPECTATIONS MODEL OF THE VOLTAGE EFFECT

Omar Al-Ubaydli
Chien-Yu Lai
John A. List

Working Paper 30850
http://www.nber.org/papers/w30850

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2023

We wish to thank Tigran Melkonyan and Rabee Tourky for very helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w30850

A Simple Rational Expectations Model of the Voltage Effect
Omar Al-Ubaydli, Chien-Yu Lai, and John A. List
NBER Working Paper No. 30850
January 2023
JEL No. C93,D61,D90

ABSTRACT

The "voltage effect" is defined as the tendency for a program's efficacy to change when it is scaled up, which in most cases results in the absolute size of a program's treatment effects to diminish when the program is scaled. Understanding the scaling problem and taking steps to diminish voltage drops are important because if left unaddressed, the scaling problem can weaken the public's faith in science, and it can lead to a misallocation of public resources. There exists a growing literature illustrating the prevalence of the scaling problem, explaining its causes, and proposing countermeasures. This paper adds to the literature by providing a simple model of the scaling problem that is consistent with rational expectations by the key stakeholders. Our model highlights that asymmetric information is a key contributor to the voltage effect.

Omar Al-Ubaydli
Department of Economics and
Mercatus Center
George Mason University
omar@omar.ec

Chien-Yu Lai
The Department of Economics
University of Chicago
E. 59th Street
Chicago, IL 60637
and Australian National University
cylai@uchicago.edu

John A. List
Department of Economics
University of Chicago
1126 East 59th
Chicago, IL 60637
and Australian National University
and also NBER
jlist@uchicago.edu

A SIMPLE RATIONAL EXPECTATIONS MODEL OF THE VOLTAGE EFFECT

January 2023

Omar Al-Ubaydli, Chien-Yu Lai, and John A. List[1]

ABSTRACT

The "voltage effect" is defined as the tendency for a program's efficacy to change when it is scaled up, which in most cases results in the absolute size of a program's treatment effects to diminish when the program is scaled. Understanding the scaling problem and taking steps to diminish voltage drops are important because if left unaddressed, the scaling problem can weaken the public's faith in science, and it can lead to a misallocation of public resources. There exists a growing literature illustrating the prevalence of the scaling problem, explaining its causes, and proposing countermeasures. This paper adds to the literature by providing a simple model of the scaling problem that is consistent with rational expectations by the key stakeholders. Our model highlights that asymmetric information is a key contributor to the voltage effect.

1. INTRODUCTION

Grounding government policy in scientific methods has become mainstream (Cartwright and Hardie, 2012), and this partially reflects the considerable advances in the science of public administration that have occurred during the postwar era (Heinrich, 2007). In democracies in particular, election manifestos and government white papers regularly cite academic papers as bases for policy proposals, or as reasons for rejecting alternatives. Moreover, the scope of the infusion of science has expanded considerably, with scholarly contributions in the social sciences playing a much larger role than before in the genesis of government policy.

In parallel to these developments, the academic disciplines informing government policy have themselves been evolving as part of the natural process of scientific advancement. In the case of economics, there has been a large growth in the contribution of experimental methods to the literature in general, and to the policy-relevant elements of the literature in particular. One byproduct of this manifestation has been the establishment of behavioral insights units in many governments across the world (Hallsworth and Kirkman, 2020). These organizations are explicitly

[1] Affiliations: Al-Ubaydli: Bahrain Center for Strategic, International and Energy Studies (Derasat), and George Mason University; Lai: University of Chicago and ANU; List: University of Chicago, ANU, and NBER. We wish to thank Tigran Melkonyan and Rabee Tourky for very helpful comments.

tasked with running rigorous small-scale experiments on matters relating to human behavior, and then using the results to make recommendations for large-scale policies.

For example, Hallsworth et al. (2017) investigated the effect of experimentally varying the wording of official reminder letters on the payment rates for overdue taxes in the UK. While the basic sample was large (200,000 participants), the study's goal was to inform the design of future official letters at the scale of the entire UK population, and to other countries looking to benefit from the experiment.

The fundamental driving force behind the increasing use of small-scale experiments in policy has been a desire to improve the quality of policymaking by ensuring that governments implement policies that actually work—and avoid ones that do not. This laudable goal has, however, been undermined by a phenomenon known as the "scaling problem" or the "voltage effect" (Al-Ubaydli et al., 2017a; Al-Ubaydli et al. 2017b; Banerjee et al., 2017; Muralidharan and Niehaus, 2017; List, 2022), which is defined as the propensity for the absolute size of an intervention's treatment effect to change when the intervention is scaled up (or, more generally, for the benefit-cost profile to change at scale).

The Head Start early childhood intervention illustrates an example of a voltage drop. A small-scale experiment based on home visits was initially implemented, yielding substantial positive treatment effects in a range of child and parental outcomes (Paulsell et al., 2010). However, when scaled, the treatment effects shrunk due to a series of factors including diminished quality of home visits (especially for at risk families) and increased attrition (Raikes et al., 2006; Roggman et al., 2008). List (2022) provides several further empirical examples of both voltage drops and voltage gains.

Describing the scaling problem as a grave threat to evidence-based policy and the public's trust in policymaking is not hyperbole. Ordinary people are cognizant of the failure of an evidence-based policy, and they update their beliefs accordingly. As an illustration, during the Covid-19 pandemic, the US government's scientists vacillated regarding the issue of wearing masks. While the changing recommendations might not have been caused by the scaling problem, the gyrations contributed to a significant decline in the public's trust in science.

This was reflected in an op-ed by the Chief Executive Officer of the American Association for the Advancement of Science calling for more careful science and better communication to maintain the public's trust (Parikh, 2021). Moreover, a 2022 poll by Pew found that from November 2020 to December 2021, the percentage of US adults with "a great deal" of confidence in medical scientists declined from 40% to 29%, while the percentage with "not too much/no confidence at all" rose from 14% to 22% (Kennedy et al., 2022). These doubts directly contributed to a decrease in the efficacy of evidence-based public health policies such as vaccines, as the population became more suspicious and less compliant.

The burgeoning aforementioned literature on the scaling problem has focused on three primary elements: demonstrating occurrences of the scaling problem (O'Donnell, 2008), as in the Early

Head Start mentioned above; explaining the causes of the scaling problem (Dusenbury et al., 2003); and proposing ex ante and ex post countermeasures that help decrease its incidence (Al-Ubaydli et al., 2021). From an economics perspective, one missing element in the above research program has been a rational expectations model of the microeconomic behavior that generates the scaling problem. Many of the actions that scholars have argued lead to the scaling problem are nominally inconsistent with a rational expectations equilibrium, suggesting that they might only account for transient instances of the phenomenon.

For example, Young et al. (2008) build a narrative model of the problem of publication bias (the tendency for journals to publish false positives), which Al-Ubaydli et al. (2017b) cite as a potential cause of the scaling problem. However, the narrative model involves both editors and the general scientific community experiencing persistent systematic errors in their assessment of the accuracy of the findings of a scientific paper: a study that reports a large treatment effect is taken at face value, and professional scientists fail to make a Bayesian adjustment that would correct for such bias. Part of the reason for the absence of rational expectations models is that the scaling problem literature is largely interdisciplinary, and the intersection of methods used is largely in the domain of empirical techniques rather than game-theoretic tools.

This paper's primary contribution is the provision of a simple, yet general, game theoretic model of the scaling problem wherein the scaling problem arises in a rational expectations equilibrium. In our model, scientists have private information regarding the size of the treatment effect of a proposed policy intervention. They report their private information to the public, including policymakers, who then make an assessment of whether to implement the proposed intervention at scale. We show how several distortions – most notably the material benefits accruing to the scientist should the program be implemented – create an incentive for the scientist to report an exaggerated treatment effect. Policymakers rationally mark this down somewhat, but the bias is not completely eliminated, and the scaling problem remains.

Our study represents an important addition to the literature because it enhances the empirical plausibility of the proposed mechanisms underlying the scaling problem. Moreover, it facilitates the refinement of the remedies that the literature recommends for overcoming the scaling problem. For example, were the scaling problem to be caused exclusively by persistent irrational expectations, then the countermeasure would be for people to correct rectify their beliefs; but when the scaling problem arises under rational expectations, then such a countermeasure would be of no value.

2. PREAMBLE ON THE SCALING PROBLEM

Al-Ubaydli et al. (2020) present a simple, decision-theoretic model of the scaling problem that does not feature rational expectations, but that nonetheless facilitates the classification of the different sources of the scaling effect. In their model, there are four broad mechanisms.

First, decreasing administration quality with scale: the technology of administering a dose (intervention) becomes more complex and difficult to implement at scale, leading to a lower fidelity to the procedure used in the initial small-scale experiment. This includes diseconomies of scale in recruitment and implementation as well.

Second, spillover effects that occur at scale. For example, when estimating the effect of enrolling in a CV-writing workshop on the rate at which job seekers receive new job offers, the results from a small-scale study are likely to exaggerate the findings from the population level, because the former does not take into account the fact that benefits accruing to those undergoing the training have negative externalities on other job seekers. At scale, the externality will be internalized, attenuating the average treatment effect. Of course, spill over effects can also be positive, leading to voltage gains at scale, such as the case for goods or services with network externalities.

Third, participants in the small-scale experiment being unrepresentative of those in the general population. In particular, the scientist conducting the small-scale study has an incentive to select people who will exhibit a larger treatment effect than the general population. This is because such individuals are cheaper to recruit and less likely to drop out of an experiment, as in the case of a medical treatment. This is also because the experiment will require fewer observations based on a power calculation and will more likely result in a peer-reviewed publication for the scientist, since editors and the scientific community in general have a preference for positive results. Finally, if the government is more likely to implement a program that showed great efficacy at small scale, and doing a large-scale rollout provides the scientist with both consulting earnings and prestige, then this creates an incentive for the scientist to handpick the participant pool and use statistical methods in favor of demonstrating a large treatment effect.

Fourth, independently of any sample manipulation by the scientist, the file drawer effect (Young et al., 2008) generates a scaling problem, whereby the tendency for editors to reject papers with negative findings and accept ones with positive findings creates an upward bias in the reported treatment effects of published studies underlain by a higher probability of type I error. All else equal, voltage drops are a result.

Notably, the first two mechanisms (diseconomies of scale, negative spillover effects) are technological and do not reflect any sort of strategic behavior by actors, and so there is no need for a rational expectations model for these features of the model. For a full model of the fourth mechanism (publication bias), see Maniadis et al. (2014).

This paper focuses on the third mechanism: strategic behavior by scientists based on exploiting their private information regarding various aspects of the small-scale study that informs the population-level rollout. The private information could reflect the nature of the participant pool, the number of times the study was actually run versus the reported number of trials, the statistical tests used, the observations classified as “outliers” and dropped, and so on.

As we will show, if scientists care sufficiently about their long-term reputation, then they will conduct small-scale studies in a manner that reflects the expected findings at scale, and there will be no scaling problem. However, if they are sufficiently short-sighted, then they will exploit the asymmetric information to manipulate the small-scale study in a manner that generates voltage drops at scale. The presence of private information regarding both the nature of the study that the scientist uses *and* their degree of short-sightedness together ensures that a rational policymaker cannot accurately infer the likelihood of an inflated reported treatment effect, generating a scaling effect that is consistent with rational expectations.

3. MODEL SETUP

3.1. PLAYERS

The game has two players (beyond nature): the government, which is considering whether to implement a program; and a researcher, who has access to more accurate information than the government on the program's effectiveness, and communicates that information to the government.

3.2. STRATEGY SPACE

Nature moves first. It determines the program's treatment effect, which is a random variable $T \in \{0, \tau\}$, where $\tau > 0$ and $\Pr(T = \tau) = 1 - \Pr(T = 0) = p \in (0,1)$. The realized value of T, denoted t, is unobservable to the government but is observable to the researcher. The researcher also privately draws their cost of dishonesty (see below).

The researcher observes t through a combination of their own insights and research. The researcher chooses $\hat{t} \in \{0, \tau\}$, which is the value of t that they report to the government. The government observes $\hat{t}$ and chooses $D(\hat{t}) \in \{0,1\}$, where $D = 1$ denotes implementation of the program, and $D = 0$ denotes non-implementation.

3.3. PAYOFFS

If the researcher reports $\hat{t}$ and the government chooses to implement the program ($D = 1$), the researcher earns a gross implementation-related payoff of:

$$\pi_R = \alpha_R + \pi_R^{\beta}(\tilde{p}(\hat{t})\tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R)$$

$$\frac{\partial \pi_R^{\beta}}{\partial \tilde{p}(\hat{t})\tau} > 0, \frac{\partial \pi_R^{\beta}}{\partial \beta_R} > 0, \frac{\partial \pi_R^{\gamma}}{\partial \tau} > 0, \frac{\partial \pi_R^{\gamma}}{\partial \gamma_R} > 0$$

$$\frac{\partial^2 \pi_R^{\beta}}{\partial \tilde{p}(\hat{t}) \tau \partial \beta_R} \geq 0, \frac{\partial^2 \pi_R^{\gamma}}{\partial \tau \partial \gamma_R} \geq 0$$

$$\pi_R^{\beta} > 0, \pi_R^{\gamma} \geq 0$$

Three terms in can $\alpha_R + \pi_R^{\beta}(\tilde{p}(\hat{t})\tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R)$ be interpreted as three types of payoff from implementation. α_R is the fixed payment. $\pi_R^{\beta}(\tilde{p}(\hat{t})\tau; \beta_R)$ is the payment from the expected treatment effect, and we could take this as from how reliable the report is from government's viewpoint. $\pi_R^{\gamma}(\tau; \gamma_R)$ is the payment from implement an effective policy. β_R and γ_R are how important the expected treatment effect and effective treatment effect are for generating researcher's payoff from implementation. Where β_R is a strictly positive constant, α_R and γ_R are non-negative constants, and $\tilde{p}(\hat{t})$ is the government's posterior probability of $T = \tau$, in light of the researcher's report $\hat{t}$. We denote that lager β_R and γ_R reflect lager marginal payoff from expected treatment effect and effective treatment effect. This payoff reflects a combination of prestige, status, and outside payments for the researcher. If the government does not implement the program $(D = 0)$, the researcher earns a gross implementation-related payoff of 0.

A researcher who lies $(\hat{t} \neq t)$ incurs a reporting loss of $c \geq 0$, which is the realized value of a random variable C with probability density function $f_C(x)$. The cost of misreporting is heterogenous across researchers, so we assume the cost variable, c, is a random draw from a density function. This reflects the long-term cost of reporting inaccurate results, which emerges when other researchers conduct follow-up work that includes replication of the original work. The cost may be paid in the future with some probability, but we could make it a discounted present value, and incorporate a heterogenous discount factor into the density function. We assume the cumulative distribution function of the cost variable, $F_C(x)$, is differentiable. Let $\bar{c} > 0$ be the largest value that C can take for which $f_C(x) > 0$. The realized value c is the researcher's private information, meaning that it is unobservable from the government's perspective. The random variable C has the moments $E(C) = \mu_C$ and $Var(C) = \sigma_C^2$.

Therefore, the researcher's net payoff u_R is:

$$u_R(t, \hat{t}, D, \tilde{p}, c) = \pi_R D - \frac{c}{\tau^2}(t - \hat{t})^2 = \left[\alpha_R + \pi_R^{\beta}(\tilde{p}(\hat{t})\tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R)\right] D - \frac{c}{\tau^2}(t - \hat{t})^2$$

The researcher has two types of information. One is the program's real treatment effect, and the other is their own cost variable. Then, the researcher decides how to make report for the program. If the researcher reports $\hat{t}$ and the government chooses to implement the program $(D = 1)$, the government earns a gross implementation-related payoff of:

$$\pi_G = \pi_G^{\beta}(\tilde{p}(\hat{t})\tau; \beta_G)$$

$$\frac{\partial \pi_G^\beta}{\partial \tilde{p}(\hat{t})\tau} > 0, \frac{\partial \pi_G^\beta}{\partial \beta_G} > 0, \frac{\partial^2 \pi_G^\beta}{\partial \tilde{p}(\hat{t})\tau \partial \beta_G} \geq 0$$

$$\tilde{p}(\hat{t})\tau > 0 \Rightarrow \pi_G^\beta > 0;\ \tilde{p}(\hat{t})\tau = 0 \Rightarrow \pi_G^\beta = 0$$

The government cares about how effective the program is. After receiving the researcher's report, the government updates the expected treatment effect of the policy and makes decisions according to this expected treatment effect. β_G is a strictly positive constant that amplifies the effect of the treatment effect on the government's payoff. The government's payoff reflects the benefits stemming from the underlying treatment effect. Implementing also costs the government a fixed amount $K > 0$.

Therefore, the government's net payoff u_G is:

$$u_G(D, \tilde{p}) = (\pi_G - K)D = \left(\pi_G^\beta(\tilde{p}(\hat{t})\tau; \beta_G) - K\right)D$$

$$D \in \{0,1\}$$

Since the government's prior for $t = \tau$ is p, we also assume that $K > K' = \pi_G^\beta(p\tau; \beta_G)$, meaning that based purely on the prior, the government never implements the program.

$$K > K' = \pi_G^\beta(p\tau; \beta_G)$$

4. EQUILIBRIUM

We now determine the unique, pure strategy Bayesian Nash equilibrium of this game. In this equilibrium, researchers with a sufficiently high cost will always truthfully report what they know about t; whereas those with a low cost will report truthfully only when $t = \tau$, while choosing to exaggerate the treatment effect ($\hat{t} > t$) when $t = 0$.

Proposition 1: If K is sufficiently small ($K < K''$) and $\bar{c} > \alpha_R + \pi_R^{\beta}(p\tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R)$ then there exists a unique pure strategy Bayesian Nash equilibrium with the following properties:

$$\hat{t}(c,t) = \begin{cases} t \; if \; c \geq c^* \\ \tau \; if \; c < c^* \end{cases}, D(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ 1 \; if \; \hat{t} = \tau \end{cases}, \tilde{p}(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ p^* \; if \; \hat{t} = \tau \end{cases}$$

Where:

$$p^* = \frac{p}{p + (1-p)F_C(c^*)} \geq p$$

$$c^* = \alpha_R + \pi_R^{\beta}(p^*\tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R) > 0$$

$$K'' = \pi_G^{\beta}(p^*\tau; \beta_G) > 0$$

Proof: We start by demonstrating the proposed equilibrium and prove uniqueness further below. Assume that the researcher is playing:

$$\hat{t}(c,t) = \begin{cases} t \; if \; c \geq c^* \\ \tau \; if \; c < c^* \end{cases}$$

According to this strategy, the researcher reports $\hat{t} = 0$ only if $t = 0$, and so $\tilde{p}(0) = 0$. In contrast, the researcher reports $\hat{t} = \tau$ either when $t = \tau$ or when $t = 0, c < c^*$. Therefore, using Bayes' rule, $\tilde{p}(\tau) = p^*$ is:

$$p^* = \frac{\Pr[\hat{t} = \tau | t = \tau] \Pr(t = \tau)}{\Pr[\hat{t} = \tau | t = 0] \Pr(t = 0) + \Pr[\hat{t} = \tau | t = \tau] \Pr(t = \tau)} = \frac{p}{(1-p)F_C(c^*) + p}$$

Thus, the proposed beliefs are correct. Given these beliefs, should the researcher report $\hat{t} = 0$ (given the rational belief that $t = 0$), implementing the program creates a strictly negative payoff of $-K$, compared to zero from not implementing:

$$u_G(1,0) - u_G(0,0) = -K < 0$$

Thus, $D(0) = 0$ is rational. Should the researcher report $\hat{t} = \tau$, then implementing the program yields the following payoff:

$$u_G(1, p^*) = \pi_G^{\beta}(p^*\tau; \beta_G) - K = K'' - K > u_G(0, p^*) = 0$$

Thus $D(\tau) = 1$ is also rational, completing our demonstration of the rationality of the government's behavior. To examine the researcher's behavior, assume that the government is playing the strategy/beliefs:

$$D(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ 1 \; if \; \hat{t} = \tau \end{cases}, \tilde{p}(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ p^* \; if \; \hat{t} = \tau \end{cases}$$

Then $u_R(t=\tau,\hat{t}=\tau,D,\tilde{p},c) > u_R(t=\tau,\hat{t}=0,D,\tilde{p},c)$, and so researchers will never underreport a high treatment effect, because they forgo the implementation benefit, and they incur the cost of dishonesty.

In contrast, if $t=0$, we have that:

$$u_R(t=0,\hat{t}=0,D,\tilde{p},c) - u_R(t=0,\hat{t}=\tau,D,\tilde{p},c)$$

$$= 0 - \left(\alpha_R + \pi_R^{\beta}(p^*\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R) - c\right) \geq 0$$

$$\Leftrightarrow\ c \geq \alpha_R + \pi_R^{\beta}(p^*\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R)$$

Thus, for sufficiently high c (for researchers who are sufficiently averse to dishonesty), the researcher will truthfully report the treatment effect $(\hat{t}=0)$; otherwise, the researcher will exaggerate the treatment effect $(\hat{t}=\tau)$.

The switching point that determines whether a researcher will be dishonest when $t=0$ is $c^* = \alpha_R + \pi_R^{\beta}(p^*\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R)$. Thus, c^* is defined by the equation:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^{\gamma}(\tau;\gamma_R)$$

There exists a unique solution to this equation. To see why, define the left-hand side and right-hand side of the equation:

$$LHS(c) = c, RHS(c) = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c)};\beta_R\right) + \pi_R^{\gamma}(\tau;\gamma_R)$$

Both are differentiable functions for $c \geq 0$. Then:

$$\lim_{c\to 0} LHS(c) = 0, \lim_{c\to\infty} LHS(c) = \infty$$

$$\lim_{c\to 0} RHS(c) = \alpha_R + \pi_R^{\beta}(\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R) > 0$$

$$\lim_{c\to\infty} RHS(c) = \alpha_R + \pi_R^{\beta}(p\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R) < \infty$$

$$\alpha_R + \pi_R^{\beta}(\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R) > \alpha_R + \pi_R^{\beta}(p\tau;\beta_R) + \pi_R^{\gamma}(\tau;\gamma_R) > 0$$

$$LHS'(c) = 1 > 0\ \forall\ c \geq 0$$

$$RHS'(c) = \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p+(1-p)F_C(c)}\right]}\left[-\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right] \leq 0\ \forall\ c \geq 0$$

$$\Rightarrow \exists!\, c \geq 0 : LHS(c) = RHS(c)$$

To demonstrate that this is the unique pure strategy equilibrium, we note that the government has two actions $(D = 0, D = 1)$ at two nodes $(\hat{t} = 0, \hat{t} = \tau)$, meaning that it has four possible pure strategies:

$$D(\hat{t}) = 0, D(\hat{t}) = 1, D(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ 1 \; if \; \hat{t} = \tau \end{cases}, D(\hat{t}) = \begin{cases} 1 \; if \; \hat{t} = 0 \\ 0 \; if \; \hat{t} = \tau \end{cases}$$

We consider each in turn.

Case 1: If $D(\hat{t}) = 0$, then:

$$u_R = -\frac{c}{\tau^2}(t - \hat{t})^2 \Rightarrow \hat{t}(c, t) = t \Rightarrow \tilde{p}(\hat{t}) = \begin{cases} 0 \; if \; \hat{t} = 0 \\ 1 \; if \; \hat{t} = \tau \end{cases}$$

Thus, the researcher will always tell the truth, and the government always knows the real treatment effect. However, since $K < K'' = \pi_G^{\beta}(p^*\tau; \beta_G) < \pi_G^{\beta}(\tau; \beta_G)$,we have that $D(\tau) = 1$ yields a higher payoff to the government than $D(\tau) = 0$, and so this cannot be a pure strategy equilibrium.

Case 2: If $D(\hat{t}) = 1$, then because $K > K' = \pi_G^{\beta}(p\tau; \beta_G)$, it must be the case that the government's posterior probability of $t = \tau$ is larger than p under both $\hat{t} = 0$ and $\hat{t} = \tau$, which is a contradiction since these are two complementary states of the world.

Case 3: Covered above in the proof of the equilibrium.

Case 4: If $D(\tau) = 0$, then $\pi_G^{\beta}(\tilde{p}(\tau)\tau; \beta_G) < K$, and if $D(0) = 1$, then $\pi_G^{\beta}(\tilde{p}(0)\tau; \beta_G) > K$, and so $\tilde{p}(\tau) < \tilde{p}(0)$. Given this, a researcher will always truthfully report $t = 0$, i.e., $\hat{t}(c, 0) = 0 \; \forall \; c$. If $\hat{t}(c, 0) = 0 \; \forall \; c$, then $\tilde{p}(\tau) = 1$. This contradicts $D(\tau) = 0$ since, $K < K'' = \pi_G^{\beta}(p^*\tau; \beta_G) < \pi_G^{\beta}(\tau; \beta_G)$.Therefore this equilibrium cannot exist. ■c

5. COMPARATIVE STATICS

For our comparative static analysis, we assume that K remains in the interval (K', K''), meaning that we are ruling out corner solutions. We are interested in the comparative statics with respect to the following parameters:

$$\beta_R, \alpha_R, \gamma_R, \beta_G, \tau, K, p$$

We are also interested in how the equilibrium is affected by changes in the cumulative density function for the random variable C.

We begin by restating the equilibrium condition:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^{\gamma}(\tau;\gamma_R)$$

We next distinguish between $\theta \in \{\beta_R, \alpha_R, \gamma_R, \beta_G, \tau, K\}$ and p. In the case of θ, we also use the definition of p^*:

$$p^* = \frac{p}{p+(1-p)F_C(c^*)}$$

According to this equation:

$$\frac{\partial p^*}{\partial \theta} = -\left(\frac{p(1-p)f_C(c^*)}{[p+(1-p)F_C(c^*)]^2}\right)\frac{\partial c^*}{\partial \theta}$$

$$\frac{\partial c^*}{\partial \theta} > 0 \Leftrightarrow \frac{\partial p^*}{\partial \theta} < 0$$

Meaning that the sign of the effect of a change in any of the parameters $\{\beta_R, \alpha_R, \gamma_R, \beta_G, \tau, K\}$ on c^* is the opposite of the sign of the effect of a change in the same parameter on p^*.

Lemma 1.1: The effect of changes in β_R:

$$\frac{\partial c^*}{\partial \beta_R} > 0, \frac{\partial p^*}{\partial \beta_R} < 0$$

Proof: Differentiating through the equilibrium condition yields:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^{\gamma}(\tau;\gamma_R)$$

$$\Rightarrow \frac{\partial c^*}{\partial \beta_R} = \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p+(1-p)F_C(c)}\right]}\left[-\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right]\frac{\partial c^*}{\partial \beta_R} + \frac{\partial \pi_R^{\beta}}{\partial \beta_R}$$

$$\Rightarrow \frac{\partial c^*}{\partial \beta_R} = \frac{\frac{\partial \pi_R^{\beta}}{\partial \beta_R}}{1 + \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p+(1-p)F_C(c)}\right]}\left[\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right]} > 0$$

$$\frac{\partial \pi_R^{\beta}}{\partial \beta_R} > 0, \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p+(1-p)F_C(c)}\right]} > 0, p(1-p)f_C(c)\tau > 0$$

$$\Rightarrow \frac{\partial c^*}{\partial \beta_R} > 0 \Rightarrow \frac{\partial p^*}{\partial \beta_R} < 0 \blacksquare$$

In other words, increasing the benefits that the researcher receives from project implementation leads to a higher cost threshold for misrepresenting low actual treatment effects, which means a higher rate of deception, and hence a lower posterior when a researcher reports $\hat{t} = \tau$. The relative sizes between the researcher's payoff and their cost of lying matter for the researcher's reporting decision. Increasing β_R is thus similar to decreasing the cost of lying, thereby inducing more lying.

Lemma 1.2: The effect of changes in α_R:

$$\frac{\partial c^*}{\partial \alpha_R} > 0, \frac{\partial p^*}{\partial \alpha_R} < 0$$

Proof: Differentiating through the equilibrium condition yields:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p + (1-p)F_C(c^*)}; \beta_R\right) + \pi_R^{\gamma}(\tau; \gamma_R)$$

$$\Rightarrow \frac{\partial c^*}{\partial \alpha_R} = \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[-\frac{p(1-p)f_C(c)\tau}{[p + (1-p)F_C(c)]^2}\right]\frac{\partial c^*}{\partial \alpha_R} + 1$$

$$\Rightarrow \frac{\partial c^*}{\partial \alpha_R} = \frac{1}{1 + \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[\frac{p(1-p)f_C(c)\tau}{[p + (1-p)F_C(c)]^2}\right]} > 0$$

$$\Rightarrow \frac{\partial c^*}{\partial \alpha_R} > 0 \Rightarrow \frac{\partial p^*}{\partial \alpha_R} < 0 \blacksquare$$

Lemma 1.3: The effect of changes in γ_R:

$$\frac{\partial c^*}{\partial \gamma_R} > 0, \frac{\partial p^*}{\partial \gamma_R} < 0$$

Proof: Differentiating through the equilibrium condition yields:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p + (1-p)F_C(c^*)}; \beta_R\right) + \pi_R^{\gamma}(\tau; \gamma_R)$$

$$\Rightarrow \frac{\partial c^*}{\partial \gamma_R} = \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[\frac{p(1-p)f_C(c)\tau}{[p + (1-p)F_C(c)]^2}\right]\frac{\partial c^*}{\partial \gamma_R} + \frac{\partial \pi_R^{\gamma}}{\partial \gamma_R}$$

$$\Rightarrow \frac{\partial c^*}{\partial \gamma_R} = \frac{\frac{\partial \pi_R^{\gamma}}{\partial \gamma_R}}{1 + \frac{\partial \pi_R^{\beta}}{\partial \left[\frac{p\tau}{p+(1-p)F_C(c^*)}\right]}\left[\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right]} > 0$$

$$\Rightarrow \frac{\partial c^*}{\partial \gamma_R} > 0 \Rightarrow \frac{\partial p^*}{\partial \gamma_R} < 0 \blacksquare$$

Lemma 2: The effect of changes in τ:

$$\frac{\partial c^*}{\partial \tau} > 0, \frac{\partial p^*}{\partial \tau} < 0$$

Using an equivalent method:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c^*)}; \beta_R\right) + \pi_R^{\gamma}(\tau; \gamma_R)$$

$$\Rightarrow \frac{\partial c^*}{\partial \tau} = \frac{\partial \pi_R^{\beta}}{\partial \left[\frac{p\tau}{p+(1-p)F_C(c^*)}\right]}\left[-\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right]\frac{\partial c^*}{\partial \tau} + \frac{\partial \pi_R^{\beta}}{\partial \left[\frac{p\tau}{p+(1-p)F_C(c^*)}\right]}\left[\frac{p}{p+(1-p)F_C(c^*)}\right] + \frac{\partial \pi_R^{\gamma}}{\partial \tau}$$

$$\Rightarrow \frac{\partial c^*}{\partial \tau} = \frac{\frac{\partial \pi_R^{\beta}}{\partial \left[\frac{p\tau}{p+(1-p)F_C(c^*)}\right]}\left[\frac{p}{p+(1-p)F_C(c^*)}\right] + \frac{\partial \pi_R^{\gamma}}{\partial \tau}}{1 + \frac{\partial \pi_R^{\beta}}{\partial \left[\frac{p\tau}{p+(1-p)F_C(c^*)}\right]}\left[\frac{p(1-p)f_C(c)\tau}{[p+(1-p)F_C(c)]^2}\right]} > 0$$

$$\Rightarrow \frac{\partial c^*}{\partial \tau} > 0 \Rightarrow \frac{\partial p^*}{\partial \tau} < 0 \blacksquare$$

Thus, the effect of increasing the size of the large treatment effect is analogous to increasing the researcher's reward in the event of a project being implemented.

Lemma 3: The effect of changes in β_G and K:

$$\frac{\partial c^*}{\partial \beta_G} = \frac{\partial c^*}{\partial K} = \frac{\partial p^*}{\partial \beta_G} = \frac{\partial p^*}{\partial K} = 0$$

Proof: Neither parameter appears in the equilibrium-defining equations. ■

Therefore, both the government's benefit from the project and the government's implementation cost have no impact on equilibrium strategies or beliefs, beyond the effect associated with allowing the condition $K' < K < K''$ to bind.

Lemma 4: The effect of changes in p:

$$\frac{\partial c^*}{\partial p} > 0, \frac{\partial p^*}{\partial p} > 0$$

Proof: Differentiating through the equilibrium condition yields:

$$c^* = \alpha_R + \pi_R^{\beta}\left(\frac{p\tau}{p + (1-p)F_C(c^*)}; \beta_R\right) + \pi_R^{\gamma}(\tau; \gamma_R)$$

$$\Rightarrow \frac{\partial c^*}{\partial p} = \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[-\frac{p(1-p)f_C(c^*)\tau}{[p + (1-p)F_C(c^*)]^2}\right]\frac{\partial c^*}{\partial p} + \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[\frac{F_C(c^*)\tau}{[p + (1-p)F_C(c^*)]^2}\right]$$

$$\Rightarrow \frac{\partial c^*}{\partial p} = \frac{\frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[\frac{F_C(c^*)\tau}{[p + (1-p)F_C(c^*)]^2}\right]}{1 + \frac{\partial \pi_R^{\beta}}{\partial\left[\frac{p\tau}{p + (1-p)F_C(c^*)}\right]}\left[\frac{p(1-p)f_C(c^*)\tau}{[p + (1-p)F_C(c^*)]^2}\right]} > 0$$

$$c^* = \alpha_R + \pi_R^{\beta}(p^* \cdot \tau; \beta_R) + \pi_R^{\gamma}(\tau; \gamma_R)$$

Higher c^* implies higher p^*, thus:

$$\frac{\partial p^*}{\partial p} > 0 \blacksquare$$

Therefore, increasing p leads to an increase in the effectiveness of a lie, ceteris paribus, which leads to more lying. The effect on the posterior probability of a treatment effect of τ given that the researcher reports $\hat{t} = \tau$ is the sum of too opposing changes: on the one hand, there is an increase in the probability of the underlying outcome, while alternatively, there is an increase in deception. The net effect is always positive.

Finally, we examine the case of changing the CDF of the cost of lying. Let $G_C(x)$ be an alternative cumulative density function (CDF) for C which first-order stochastically dominates the CDF

$F_C(x)$, meaning that $G_C(x) \leq F_C(x)\ \forall\ x$. Let c^{**} be the equilibrium switching point associated with the CDF $G_C(x)$.

Lemma 5: If $G_C(x)$ first-order stochastically dominates $F_C(x)$, then:

$$c^{**} \geq c^*, G_C(c^{**}) \leq F_C(c^*)$$

Proof:

$$c^* = \alpha_R + \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^\gamma(\tau;\gamma_R)$$

$$c^* - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) = \alpha_R + \pi_R^\gamma(\tau;\gamma_R) = c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right)$$

$$\Rightarrow c^* - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) = c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right)$$

$$\Rightarrow c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right) < c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^{**})};\beta_R\right)$$

$$\Rightarrow c^* - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) < c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^{**})};\beta_R\right)$$

$$\Rightarrow c^* - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) - \alpha_R - \pi_R^\gamma(\tau;\gamma_R)$$
$$< c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^{**})};\beta_R\right) - \alpha_R - \pi_R^\gamma(\tau;\gamma_R)$$

$$\Rightarrow 0 < c^{**} - \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^{**})};\beta_R\right) - \alpha_R - \pi_R^\gamma(\tau;\gamma_R)$$

Thus, we know

$$c^{**} > c^*$$

$$\Rightarrow \alpha_R + \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^\gamma(\tau;\gamma_R) = c^* \leq c^{**}$$
$$= \alpha_R + \pi_R^\beta\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right) + \pi_R^\gamma(\tau;\gamma_R)$$

$$\Rightarrow \alpha_R + \pi_R^\beta\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) + \pi_R^\gamma(\tau;\gamma_R)$$
$$\leq \alpha_R + \pi_R^\beta\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right) + \pi_R^\gamma(\tau;\gamma_R)$$

$$\Rightarrow \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)F_C(c^*)};\beta_R\right) \leq \pi_R^{\beta}\left(\frac{p\tau}{p+(1-p)G_C(c^{**})};\beta_R\right)$$

$$\Rightarrow F_C(c^*) \geq G_C(c^{**}) \blacksquare$$

In other words, raising the cost of dishonesty leads to lower levels of strategic dishonesty in equilibrium, and hence a lower incidence of voltage drops, or an attenuation of the scaling problem.

6. DISCUSSION

Our simple model demonstrates that asymmetric information can contribute to the scaling problem by giving scientists an incentive to surreptitiously manipulate elements of their small-scale studies in an attempt to inflate reported treatment effects. The academy's first line of defense is the importance placed on replication, but scientists differ in their long-sightedness, which is itself private information, and so this safeguard can fail to expunge the strategic sources of the scaling problem. The comparative statics associated with the model, reflected in the series of lemmas, suggests a series of theoretically-motivated countermeasures to this strategic form of the scaling problem.

First, consider decreasing the size of the reward that scholars receive when their small-scale study is rolled out to a larger population, whether it is the material consulting fees, or the prestige; because this reward creates the incentive to deceive. Admittedly, there are important general equilibrium effects that our model has not accounted for, such as the positive role that such incentives play in motivating policy-relevant research and good quality scholarship. However, at the margin, as can be seen in high status cases of scientific fraud motivated by the desire for professional rewards, there may be benefits from decreasing the size of such rewards, and for relying more on scientists' intrinsic motivation for their scholarly work.

Second, take steps to increase the cost of dishonesty. As discussed in Maniadis et al. (2014), this can be done ex post by allocating a greater volume of the academy's resources to replication and affording such replications more space in leading journals. It can also be done by imposing more stringent ex post data sharing requirements that help uncover the deployment of dubious statistical tests; as well as by convening investigation teams that conduct systematic audits on a randomly selected proportion of published studies.

There are ex ante measures, too, such as pre-registration (Gehlbach and Robinson, 2018), and improving efforts at educating aspiring scientists about the nature of scientific misconduct. Such measures can establish a more honest culture. For example, in clinical work, there are many measures taken that involve giving physicians material incentives to act in the patient's interests; but these extrinsic measures are complemented by the Hippocratic oath, which has been shown to play a substantive role in promoting ethical behavior (Askitopoulou and Vgontzas, 2018). We are

not aware of the existence of such an oath in the social sciences, and so perhaps it is time to imbue a broader sense of responsibility among scientists toward their communities.

7. CONCLUSION

In principle, grounding policy in rigorous science is highly desirable. In practice, if the science is shaky, or if it is susceptible to self-serving manipulation by the scholars that runs counter to the public interest, then grounding policy in rigorous science becomes dangerous. The scaling problem is a manifestation of this perilous dilemma: we want our large-scale interventions to be informed by sound scientific evidence gathered at a smaller scale, but we want those small-scale experiments to be consistent in the prescriptions that they provide. As the vaccine rollout in the US demonstrated, when the public's faith in science is shaken, there can be profound adverse consequences for all of society.

This paper adds to the growing literature on the scaling problem (voltage effect) by demonstrating that some of its alleged causes – strategic manipulation of small-scale studies by scientists – are consistent with equilibrium behavior and rational expectations in a game-theoretic environment. This facilitates the development of countermeasures to voltage drops, as explanations that are inconsistent with rational expectations arguably reflect transient and self-rectifying phenomena that might not merit the attention of policymakers.

For the purposes of parsimonious exposition, the environment we construct is quite simple. For example, we have a dichotomous rather than continuous treatment effect, and the payoff functions are quasi-linear in certain parameters. We hope that future research can generalize this model to explore the robustness of the conclusions to broader settings and examine drivers of voltage effects that go beyond asymmetric information.

REFERENCES

Al-Ubaydli, O., List, J.A., LoRe, D. and Suskind, D., 2017a. Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, *31*(4), pp.125-44.

Al-Ubaydli, O., List, J.A. and Suskind, D.L., 2017b. What can we learn from experiments? Understanding the threats to the scalability of experimental results. *American Economic Review*, *107*(5), pp.282-86.

Al-Ubaydli, O., List, J.A. and Suskind, D., 2020. 2017 Klein Lecture: The Science Of Using Science: Toward An Understanding Of The Threats To Scalability. *International Economic Review*, *61*(4), pp.1387-1409.

Al-Ubaydli, O., Lee, M.S., List, J.A., Mackevicius, C.L. and Suskind, D., 2021. How can experiments play a greater role in public policy? Twelve proposals from an economic model of scaling. *Behavioural public policy*, *5*(1), pp.2-49.

Askitopoulou, H. and Vgontzas, A.N., 2018. The relevance of the Hippocratic Oath to the ethical and moral values of contemporary medicine. Part II: interpretation of the Hippocratic Oath—today's perspective. *European Spine Journal*, *27*(7), pp.1491-1500.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M. and Walton, M., 2017. From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, *31*(4), pp.73-102.

Cartwright, N. and Hardie, J., 2012. *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.

Dusenbury, L., Brannigan, R., Falco, M. and Hansen, W.B., 2003. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health education research*, *18*(2), pp.237-256.

Gehlbach, H. and Robinson, C.D., 2018. Mitigating illusory results through preregistration in education. *Journal of Research on Educational Effectiveness*, *11*(2), pp.296-315.

Hallsworth, Michael, and Elspeth Kirkman. *Behavioral insights*. MIT Press, 2020.

Hallsworth, M., List, J.A., Metcalfe, R.D. and Vlaev, I., 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, pp.14-31.

Heinrich, C.J., 2007. Evidence-based policy and performance management: Challenges and prospects in two parallel movements. *The American Review of Public Administration*, *37*(3), pp.255-277.

Kennedy, B., Tyson, A., and Funk, C., 2022. Americans' Trust in Scientists, Other Groups Declines. *Pew Research Center Reports*, https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/ (accessed 2022/07/01).

List, J.A., 2022. The Voltage Effect: How to Make Good Ideas Great and Great Ideas Scale. Currency, Penguin Random house.

Maniadis, Z., Tufano, F. and List, J.A., 2014. One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, *104*(1), pp.277-90.

Muralidharan, K. and Niehaus, P., 2017. Experimentation at scale. *Journal of Economic Perspectives*, *31*(4), pp.103-24.

O'Donnell, C.L., 2008. Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of educational research*, *78*(1), pp.33-84.

Parikh, S., 2021. Why We Must Rebuild Trust in Science. *Trend Magazine*, Winter 2021.

Paulsell, D., Avellar, S., Martin, E.S. and Del Grosso, P., 2010. *Home visiting evidence of effectiveness review: Executive summary* (No. 5254a2ab30e146ce900220dbc4f41900). Mathematica Policy Research.

Raikes, H., Green, B.L., Atwater, J., Kisker, E., Constantine, J. and Chazan-Cohen, R., 2006. Involvement in Early Head Start home visiting services: Demographic predictors and relations to child and parent outcomes. *Early Childhood Research Quarterly*, *21*(1), pp.2-24.

Roggman, L.A., Cook, G.A., Peterson, C.A. and Raikes, H.H., 2008. Who drops out of early head start home visiting programs?. *Early Education and Development*, *19*(4), pp.574-599.

Young, N.S., Ioannidis, J.P.A. and Al-Ubaydli, O., 2008. Why current publication practices may distort science. *PLoS medicine*, *5*(10), p.e201.