Rational Expectations and
The Paradox of Policy-Relevant Natural Experiments

Gilles Chemla[a,*], Christopher A. Hennessy[b‡]

[a] Imperial College Business School, DRM/CNRS and CEPR;

[b] London Business School, CEPR and ECGI.

Received Date: January 2018, Received in Revised Form February 2019

Abstract

Policy experiments using large microeconomic datasets have recently gained ground in macroeconomics. Imposing rational expectations, we examine robustness of evidence derived from ideal natural experiments applied to atomistic agents in dynamic settings. Paradoxically, once experimental evidence is viewed as sufficiently clean to use, it then becomes contaminated by *ex post endogeneity*: Measured responses depend upon priors and the objective function into which evidence is fed. Moreover, agents' policy beliefs become endogenously correlated with their causal parameters, severely clouding inference, e.g. sign reversals and non-invertibility may obtain. Treatment-control differences are contaminated for non-quadratic adjustment costs. Constructively, we illustrate how inference can be corrected accounting for feedback and highlight factors mitigating contamination.

Keywords: Natural Experiments, Rational Expectations, Causality, Policy.

JEL classification: E6, G18, G28, G38, H00, O2

---

1. Introduction

In recent years, there have been calls for macroeconomists to adopt the methodology of applied microeconometricians. Romer and Romer (2014) argue, "In microeconomic settings,

---

[*] Corresponding author. Imperial College Business School, South Kensington campus, London SW7 2AZ, United Kingdom. E-mail: g.chemla@imperial.ac.uk. Tel: 44 (0) 20 7594-9161.

[†] London Business School, Regent's Park, London, NW1 4SA, U.K. E-mail: chennessy@london.edu. . Tel: 44 (0) 20 7000-8285.

[‡] We thank seminar participants at Wharton, Stanford, Berkeley, Columbia, CMU, MIT, UCLA, LBS, BC, INSEAD, IESE, Imperial, Baruch, Copenhagen, Washington-Seattle, Miami, Paris-Dauphine, Zurich, Arison, Nova-Lisbon, CEPR ESSET, and FTG. Hennessy acknowledges funding from the ERC. This paper was previously entitled Use It and Lose It: How Policy Feedback Contaminates Shock-Based Inference.

it is often possible to identify natural experiments where it is clear that differences among economic actors are not the result of confounding factors." In their influential textbook, *Mostly Harmless Econometrics*, Angrist and Pischke (2009) argue empirical evidence derived from exogenous shocks represents a credible stand-alone product: "A principle that guides our discussion is that most of the estimators in common use have a simple interpretation that is not heavily model dependent." Angrist and Pischke (2010) speak of a "credibility revolution" in microeconometrics and assert "today's macro agenda is empirically impoverished... The theory-centric macro fortress appears increasingly hard to defend."

The credibility some attribute to evidence derived from natural experiments has led to calls to make them central in policy setting. Former Minneapolis Fed President Narayana Kocherlakota (2018) argues "there has been a revolution in applied microeconometrics in the use of atheoretical statistical methods... a similar change could be of value in applied macroeconomics." Former CEA Chair Michael Greenstone (2009) goes further in calling for "persistent regulatory experimentation."

In this paper, we view seemingly-ideal experiments through the lens of rational expectations and demonstrate an important internal inconsistency. A simple example illustrates. The two primary objectives in applied microeconometric work are to demonstrate clean identification and direct policy relevance. Suppose then that an econometrician demonstrates nature forced a change in government policy or randomly assigned policy treatments. Suppose she can also demonstrate direct policy-relevance, with the experiment outcome being utilized by the government in deciding future policy.

What has gone unnoticed here is a contradiction between the econometrician's claim of clean identification and her demonstration of policy relevance. After all, in establishing policy-relevance, the econometrician has also established that the probability distribution of the policy variable is being altered by the experimental evidence. But this implies rational agents will change their behavior during the experiment. That is, expectation of endogenous evidence-based policymaking after an experiment will change what the econometrician measures during the experiment. Importantly, as is in our model, this is true even if each agent is measure zero and has no strategic motive to change behavior with the goal of influencing policy.

We consider the following economy. Measure zero corporations operating across a finite number of industries make investment decisions in light of current and expected tax rates. Industry-specific tax responsiveness parameters are privately observed, being i.i.d. draws from a known probability distribution. Econometricians want to estimate these parameters since aggregate investment is proportional to the average value. Ideal randomized evidence will arrive and econometricians will be able to exploit cross-sectional data, often argued to be an advantage of microeconomic data, e.g. Romer and Romer (2014). The evidence takes one of two forms. In an *Economy-Wide Natural Experiment* (EWNE below), all firms face a common tax rate shock during a random experiment period. In a *Controlled Natural Experiment* (CNE below), firms are randomly assigned to high or low tax rates during a random experiment period.

As shown, even with measure-zero agents, feedback from experimental evidence to the probability distribution of the policy variable post-experiment contaminates the formerly-clean evidence. However, this problem of *ex post endogeneity* vanishes if the government cannot change future policy or if the government does not view the experimental evidence as credible and ignores it. We thus have the following paradoxical situation: The clean experimental evidence is uncontaminated only if the government is unable or unwilling to use it.

We illustrate *seven novel challenges* to causal parameter inference arising from ex post endogeneity. First, rather than being stand-alone objects that are "not heavily model-dependent," policy-relevant experimental evidence must be interpreted in light of the governmental objective function into which the evidence will be fed. Second, correct causal parameter inference requires correctly stipulating agent prior beliefs regarding the probability distribution governing these same causal parameters. Third, with ex post endogeneity, the act of observation by econometricians changes behavior. Fourth, such observer effects are unequal across treatment and control groups absent strong functional form assumptions. Fifth, heterogeneous causal effect parameters generate endogenously heterogeneous policy beliefs, clouding inference. Sixth, policy feedback may cause experimental moments to become non-monotone, preventing moment inversion and parameter inference. Finally, policy feedback can result in sign bias, in addition to magnitude bias.

We do not claim bias will always be large. Rather, we argue bias will be largest if discount rates are low, so that decisions of experimental subjects are more sensitive to policy expectations. Bias will also be high if endogenous government responses are expected to occur soon after the experiment, and if government enjoys discretion to change policy by a large amount based on the experimental evidence. Conversely, biases arising from ex post endogeneity will be small when: discount rates are high; the experiment will only inform policy in the distant future; and the experimental data will have a negligible effect on the policy variable. But note, if a natural experiment meets these three conditions, it has relatively low social present value. That is, bias from ex post endogeneity is small (large) if the social value of the experimental study is relatively small (large).

The issues raised here are related to, but differ from, the econometric critique of Lucas (1976).[1] Writing for *New Palgrave Dictionary of Economics*, Ljungqvist (2008) defines the Lucas Critique as follows:

> It criticizes using estimated statistical relationships from past data to forecast effects of adopting a new policy, because the estimated coefficients are not invariant but will change along with agents' decision rules in response to a new policy.
>
> A classic example of this fallacy was the erroneous inference that a regression of

---

[1] See Linde (2001) for a discussion of empirical evidence on the Lucas Critique.

> inflation on unemployment (the Phillips curve) represented a structural trade-off for policy to exploit.

Thus, the argument of Lucas (1976) is that future regression coefficients will differ from those estimated presently if the policy rule changes. Our argument is that there will be a change in what is measured presently in light of anticipation of how experimental evidence will be used in policy setting. However, the second and most important difference is that Lucas considers exogenous policy changes. In contrast, the essential ingredient in our model is endogenous policy post-experiment. A third difference is that we utilize the heterogeneous causal parameters setup favored by microeconometricians to show that with endogenous policy, heterogeneous parameters give rise to heterogeneous policy expectations and concomitant challenges to inference. Fourth, in Lucas (1976), econometricians and their estimates are irrelevant to policy and equilibrium outcomes. In contrast, econometricians sit inside our model, with our focus being on the feedback between econometricians, their perceived credibility, and government policy. Finally, Lucas is silent on controlled trials and difference in differences estimators, while we demonstrate strong functional form requirements. These differences notwithstanding, we borrow from Lucas (1976) the idea of viewing empirical evidence through the lens of rational expectations.

Leeper and Zha (2003) show linear projections may be reliable under a latent regime-switching policy where the Lucas Critique would seem important. We consider a fiscal policy setting, where the policy variable, the corporate tax rate, is common knowledge and expectational effects are generally large under plausible parameters. Kocherlakota (2018) provides a formal rationale for theory-free regressions in a Bayesian-Nash equilibrium where the government enjoys access to an infinite data set where all possible policies occur with positive probability. We argue clean policy shocks are rare and full policy experimentation is generally infeasible. Thus, we analyze what can be learned from a single "ideal" policy experiment.

The macroeconometric literature has focused primarily on the implications of rational expectations for interpreting vector autoregressions. Sargent (1971, 1973, 1977) and Taylor (1979) showed that rational expectations imply restrictions on distributed lags. Sims (1982) and Sargent (1984) pointed to a problem in rational expectations econometrics in postulating suboptimal behavior by governments. In our model, the government uses its information optimally, substantially complicating inference.

Heckman (2000) argues agents may violate random assignment. In our economy, agents cannot do so. Heckman also emphasizes the probability limit of instrumental variables estimators can depend on the choice of instrument. In our model, there is no instrumentation. Deaton (2010) emphasizes small sample bias and biased panel selection. We consider a continuum of firms with ideal first-stage policy randomization.

Acemoglu (2010) argues general equilibrium effects, such as price changes, can limit the external validity of small-scale experiments. These effects are shut off in our model. Chassang et al (2012) consider static RCTs and show how hidden effort during an experiment can cloud inference. Their model abstracts from endogenous post-experiment policy, so the bias causes differ. Hennessy and Strebulaev (2018) analyze econometric evidence derived from an infinite sequence of natural experiments, with zero endogeneity bias.

Wolpin (2013) argues for the need to filter econometric evidence through the lens of structural models. Closer to our approach is structural corporate finance work by Gomes (2001), Alti (2003), Moyen (2005), and Hennessy and Whited (2005) who demonstrate pitfalls in inference from reduced-form investment regressions. Our paper is distinct in that it formally analyzes natural experiments and policy-feedback effects.

The rest of the paper is as follows. Section 2 presents a model of the interaction between firms, governments, and econometricians. Section 3 discusses econometric inference in settings where firms face a common economy-wide policy shock (EWNEs). Section 4 discusses controlled natural experiments (CNEs).

2. The Model

We begin by contrasting inference in two economies endowed with identical natural experiments and technologies but differing in whether the empirical evidence will be used. In the *Endogenous Policy Economy*, the evidence will be used to select an optimal policy post-experiment. In the *Exogenous Policy Economy*, the government is powerless to influence the policy variable.

2.1. Technology

The technology builds on that of Abel and Eberly (1997). Time is continuous and the horizon infinite. Agents are risk-neutral and share the discount rate $r > 0$. Tildes denote random variables and bold-type denotes vectors.

Firms are taxed on earnings before interest, taxes, depreciation, and amortization (EBITDA). The tax rate at time t is denoted τ_t. The set of politically feasible tax rates is $\tau_t \in [\underline{\tau}, \overline{\tau}]$, where $0 \leq \underline{\tau} < \overline{\tau} < 1$. The tax rate is the only policy variable.

There is a measure one continuum of anonymous firms with generic member $j \in \mathbb{J}$. Since firms are atomistic, no firm has incentive to change behavior with the goal of influencing test statistics and government policy. Each firm belongs to one of $M \geq 2$ industries of equal measure. A generic industry is denoted m.

Time and firm identifiers are omitted where obvious. The law of motion for a firm's capital stock is:

$$dk_t = (i_t - \delta k_t)dt. \tag{1}$$

In the preceding equation, i_t denotes instantaneous gross investment and $\delta \geq 0$ denotes the depreciation rate. Each firm invests optimally at each instant.

The instantaneous EBITDA flow accruing to the firm at time t is $x_t k_t$. The EBITDA factor x is a positive geometric Brownian motion with the following law of motion:

$$dx_t = \mu x_t dt + \sigma x_t dw_t. \tag{2}$$

The variable w is an independent Wiener process. The EBITDA factor x is common to all firms and thus captures macroeconomic risk.

The investment cost function for each firm in industry m is:

$$\psi(i) \equiv \gamma_m i^{\nu/(\nu-1)}, \tag{3}$$

where $\gamma_m > 0$ and $\nu \in \{2, 4, 6, ...\}$. The assumption that ν is a positive even integer follows the treatment in Abel and Eberly (1997) and ensures investment costs are real valued and convex for positive and negative investment. The parameter ν is common knowledge. However, the industry-specific investment cost parameter γ_m is known only to firms in industry m. These parameters represent realizations of i.i.d. random variables. Econometricians and the government want to infer the investment cost parameters since this will allow them to infer the tax-responsiveness of aggregate investment.

The net cash flow for a generic firm in industry m is EBITDA less taxes less investment costs:

$$(1 - \tau_t)x_t k_t - \gamma_m i_t^{\nu/(\nu-1)}. \tag{4}$$

To ensure firm value is bounded, assume

$$r > \mu\nu + \frac{1}{2}\sigma^2\nu(\nu - 1).$$

2.2. Timing

There are three stages, $S \in \{P, E, I\}$. A firm's tax rate is constant within each stage, but varies across stages. The tax rate during stage S is denoted τ_S. The Pre-Experiment Stage P is followed by the Experiment Stage E which is followed by the Implementation Stage I.

During Stage P, firms face the same tax rate $\tau_P \in [\underline{\tau}, \overline{\tau}]$. An exogenous natural experiment will arrive at time $\widetilde{t}_E$. This time is an independent random variable. The transition rate into the Experiment Stage is $\lambda_E > 0$. Thus, at any time prior to the transition, the expected remaining duration of Stage P is λ_E^{-1}.

During the Experiment Stage, nature randomly assigns a fraction θ of firms to “maximum tax status” where $\tau_E = \overline{\tau}$, and all other firms are assigned to “minimum tax status” where $\tau_E = \underline{\tau}$. In an *Economy-Wide Natural Experiment* (EWNE below), all firms face a common tax rate $\tau_E \neq \tau_P$ during the Experiment Stage. In a *Controlled Natural Experiment* (CNE below), firms are randomly assigned to treatment and control groups, one group facing the maximum and the other facing the minimum tax rate.

The Implementation Stage I will arrive at time $\widetilde{t}_I$. This date is an independent random variable given the realization of $\widetilde{t}_E$. The transition rate into the Implementation Stage is $\lambda_I > 0$. Thus, at any instant during the experiment, the expected remaining duration of Stage E is λ_I^{-1}.

At the very start of Stage I, econometricians observe some experimental evidence.[2] The evidence is such that the causal parameter vector $\boldsymbol{\gamma}$ can potentially be correctly inferred but the econometrician must understand the subtle interplay between evidence, policy, and firm-level expectations.[3] In a CNE, econometricians can look back and measure the difference between the investment of firms facing the maximum and minimum tax rates, industry-by-industry. In an EWNE, econometricians can look back and measure the jump in each industry’s investment at the start of the experiment. Since the path of x is continuous, there is no need to control for changes in macroeconomic conditions.

The following assumption is satisfied by the tax rate process $\widetilde{\tau}_{jt}$ facing arbitrary firm j

---

[2] Letting firms observe the same evidence as econometricians has no effect.

[3] Observation during Stage E adds an uninteresting waiting phase where τ_I is determined, with τ_E still in effect.

at any point in time t at which the experimental measurement may take place.

Assumption 1 (Policy Independence): $\widetilde{\tau}_{jt} \perp \{\widetilde{\boldsymbol{\gamma}}, \widetilde{x}_t\} \quad \forall \quad j \in \mathbb{J}$ and $t \in [0, \widetilde{t}_I)$.

By construction, we rule out the forms of endogeneity bias that presently occupy the attention of applied microeconometricians. First, the tax rate process is independent of the macroeconomic state at all points in time. Second, there is no Experiment Stage tax rate selection by firms or the government based on unobservables ($\boldsymbol{\gamma}$). However, one of our key points is that standard orthogonality assumptions as invoked in applied work focus on the experimental period and fail to account for correlation between unobservables and policy *post-experiment.*

During the Implementation Stage, a long-term tax rate (τ_I) will be implemented permanently. In the Exogenous Policy Economy, τ_I will be set to a technologically pre-determined value τ_I^{EX} and the government cannot change it. In the Endogenous Policy Economy, the government will infer $\boldsymbol{\gamma}$ and implement an optimal tax rate.

2.3. Endogenously Heterogeneous Policy Beliefs

Optimal investment equates marginal cost with the shadow value of installed capital, denoted q. The function i_m^* maps industry-specific q_m to optimal industry-specific investment:

$$q_m = \psi'(i_m^*) \Rightarrow i_m^*(q_m) \equiv \left(\frac{1}{\gamma_m}\right)^{\nu-1} \left(\frac{\nu-1}{\nu}\right)^{\nu-1} q_m^{\nu-1}. \tag{5}$$

Optimal investment is increasing in q, with q-sensitivity decreasing in γ_m. For example,

$$\nu = 2 \Rightarrow i_m^*(q_m) \equiv \frac{1}{2}\left(\frac{1}{\gamma_m}\right) q_m. \tag{6}$$

We now define a proxy for industry-level *investment responsiveness*:

$$\widetilde{\xi}_m \equiv \left(\frac{1}{\widetilde{\gamma}_m}\right)^{\nu-1} \Rightarrow \xi_m \equiv \left(\frac{1}{\gamma_m}\right)^{\nu-1}. \tag{7}$$

Optimal investment from equation (5) can then be expressed as:

$$i_m^*(q_m) \equiv \xi_m \left(\frac{\nu-1}{\nu}\right)^{\nu-1} q_m^{\nu-1}. \tag{8}$$

For the remainder of the paper we shall speak of the econometrician as estimating each industry's investment responsiveness parameter ξ_m, which is the realization of $\widetilde{\xi}_m$. These variables are i.i.d. draws made by nature just before the model opens. Specifically, the investment responsiveness parameter for each industry represents a draw from an interval $[\underline{\xi}, \overline{\xi}]$ on the positive real line with a strictly positive probability density f on this support and a corresponding cumulative distribution F that is twice continuously differentiable, with $F(\underline{\xi}) = 0$. Below we speak of the cumulative distribution function F as capturing *prior beliefs* (regarding the unknowns in this economy, the realized causal parameter vector $\boldsymbol{\xi}$).

During the Implementation Stage, firms share the same instantaneous EBIT factor x_t and same tax rate τ_I. Thus, firms share the same q value during this stage. This shadow value must satisfy the following equilibrium condition which demands that the instantaneous expected holding return on a unit of capital is just equal to the opportunity cost $r+\delta$. Using Ito's Lemma to compute the expected instantaneous capital gain, we have:

$$(r+\delta)q(x) = \mu x q'(x) + \frac{1}{2}\sigma^2 x^2 q''(x) + (1-\tau_I)x. \tag{9}$$

Solving, the preceding differential equation we obtain the following Implementation Stage shadow value:

$$q^I(x_t, \tau_I) = \frac{(1-\tau_I)x_t}{r+\delta-\mu}. \tag{10}$$

Aggregate investment during the Implementation Stage will be:

$$i_{AGG}(\tau_I) = \left(\frac{1}{M}\sum_{m=1}^{M} \xi_m\right) \left(\frac{\nu-1}{\nu}\right)^{\nu-1} \left(\frac{(1-\tau_I)x_t}{r+\delta-\mu}\right)^{\nu-1}. \tag{11}$$

Notice, the preceding equation implies the difference between aggregate investment across alternative Implementation Stage tax rates is proportional to the average of industry-level investment responsiveness parameter.

When the Implementation Stage is reached, the government in the Endogenous Policy Economy will permanently implement a tax rate τ_I^* to maximize its objective function Θ, where Θ is common knowledge. In particular,

$$\tau_I^*(\boldsymbol{\xi}) \in \arg\max_{\tau \in [\underline{\tau}, \overline{\tau}]} \; \Theta(\tau, \boldsymbol{\xi}) \tag{12}$$

where

$$\Theta(\tau, \boldsymbol{\xi}) \equiv \tau \times \left[\xi^* - \frac{1}{M} \sum_{m=1}^{M} \xi_m \right]. \tag{13}$$

Under this objective function the government will tax firms at the minimum (maximum) rate if the induced increase (decrease) in aggregate investment will be sufficiently large (small). In particular, if average investment responsiveness weakly exceeds the cutoff ξ^*, the government will implement the minimum (maximum) feasible tax rate.[4]

It is assumed $\xi^* \in (\underline{\xi}, \overline{\xi})$ which implies the government's policy decision can be influenced by the econometric evidence. We then consider Rational Expectations Equilibria in which the government, aided by its in-house econometrician, is able to infer $\boldsymbol{\xi}$ based on the econometric evidence at its disposal. As discussed below, in some cases the standard test statistic will not suffice for this purpose and so other statistics must be examined.

Let χ be an indicator function for average investment responsiveness weakly exceeding the government's cutoff:

$$\frac{1}{M} \sum_{m=1}^{M} \xi_m \geq \xi^* \Leftrightarrow \chi(\boldsymbol{\xi}) = 1. \tag{14}$$

The policy belief function β reflects updated assessments of the probability of the government implementing the minimum tax rate during the Implementation Stage, with the updating based upon the realized value of own-industry investment responsiveness. We have:[5]

$$\beta(\xi) = \int_{\underline{\xi}}^{\overline{\xi}} ... \int_{\underline{\xi}}^{\overline{\xi}} \int_{\underline{\xi}}^{\overline{\xi}} ... \int_{\underline{\xi}}^{\overline{\xi}} \chi(\xi_1, ..., \xi_{m-1}, \xi, \xi_{m+1}, ..., \xi_M) F(d\xi_1) ... F(d\xi_{m-1}) F(d\xi_{m+1}) ... F(d\xi_M). \tag{15}$$

---

[4] This bang-bang policy simplifies the algebra and exposition but is not essential.

[5] It is at this stage that uncertainty over the government's objective would necessitate integrating over the set of possible objectives.

Beliefs take the same functional form (β) for all industries. However, the realized value of the function's argument ξ will differ across industries with probability one.

The following lemma, which follows directly from equation (15), summarizes some important properties of the belief function.

Lemma 1 *If government can implement optimal policies in response to experimental evidence, a firm's assessment of the probability of the government implementing the minimum tax rate is non-decreasing in its own causal parameter* ξ_j*, and strictly increasing on a set of positive measure. Heterogeneous causal parameters* ($\boldsymbol{\xi}$) *result in endogenously heterogeneous policy beliefs on a set of positive measure.*

Intuitively, a firm's knowledge that its own industry is highly responsive to taxes rationally assigns a higher probability to the government choosing a low tax rate in the long-term. To illustrate, consider an economy with two industries ($M = 2$). From equation (15) it follows:

$$\beta(\xi) = 1 - F(2\xi^* - \xi) \Rightarrow \beta'(\xi) = f(2\xi^* - \xi) \geq 0. \tag{16}$$

Consistent with Lemma 1, this belief function is non-decreasing. Moreover, different realized values of the causal parameters (ξ_1, ξ_2) will generally result in endogenously heterogeneous policy beliefs. Further, the shape of the belief function is determined by prior beliefs (F), as well as the government objective function parameter ξ^*.

Continuing this example, Figure 1 plots the belief β as a function of ξ. We consider three specifications of the belief function: Baseline, High Cutoff, and Negative Priors. In the *Baseline* specification, F is the uniform distribution on $[0, 1]$ and $\xi^* = 0.50$. The *High Cutoff* specification considers the same prior beliefs F as the Baseline, but assumes $\xi^* = 0.75$. Finally, the *Negative Priors* specification considers $\xi^* = 0.75$ with F being a triangular distribution on $[0, 1]$ with mode at 0.

In Figure 1, each belief function is non-decreasing. Thus, different realizations of ξ across the industries will generally lead to different expectations. Further, changes in the government's objective function (ξ^*) and changes in prior beliefs (F) lead to endogenous changes in the belief function. Anticipating, since investment depends upon beliefs, experimental test statistics will change with changes to the government's objective function or changes to prior beliefs.

It may be tempting to argue that correct inference boils down to correctly stipulating the expected long-term tax rate. However, Lemma 1 shows there is no reason to assume agents will hold common policy expectations. Further, assuming one can make correct assumptions about each agent's belief ($\beta(\xi)$) is tantamount to assuming one knows the true value of the causal parameters ($\boldsymbol{\xi}$) one is hoping to estimate.

2.4. Firm-Level Investment

The model is formally solved in the appendix. This subsection characterizes firm investment. Throughout, upper (lower) bars denote values and policies if the current stage $\tau = \overline{\tau}$ ($\tau = \underline{\tau}$).

Equation (8) expresses investment as a function of q. Implementation Stage q was presented in equation (10). During the Experiment Stage, q as computed by a firm in industry m, must satisfy the following equilibrium condition demanding the expected holding return on capital equals the opportunity cost $r + \delta$:

$$\begin{aligned}(r+\delta)q(x) &= \mu x q'(x) + \frac{1}{2}\sigma^2 x^2 q''(x) + (1-\tau_E)x \\ &\quad + \lambda_I \left[\beta(\xi_m)\frac{(1-\underline{\tau})x}{r+\delta-\mu} + (1-\beta(\xi_m))\frac{(1-\overline{\tau})x}{r+\delta-\mu} - q(x)\right].\end{aligned} \tag{17}$$

We conjecture the preceding differential equation has a linear solution. Substituting in the conjectured linear solution and solving, one obtains:

$$q^E(x,\xi_m) = \frac{(1-\tau_E)x}{r+\delta-\mu} + \frac{\lambda_I[\tau_E - \mathbb{E}(\tau_I|\xi_m)]x}{(r+\delta-\mu)(r+\delta+\lambda_I-\mu)}. \tag{18}$$

where

$$\mathbb{E}[\tau_I|\xi_m] \equiv \beta(\xi_m)\underline{\tau} + [1-\beta(\xi_m)]\overline{\tau}. \tag{19}$$

Consider next q during the Pre-Experiment Stage. Again the expected holding return on capital equals the opportunity cost $r+\delta$:

$$\begin{aligned}(r+\delta)q(x) &= \mu x q'(x) + \frac{1}{2}\sigma^2 x^2 q''(x) + (1-\tau_P)x \\ &\quad + \lambda_E\left[\theta \overline{q}^E(x,\xi_m) + (1-\theta)\underline{q}^E(x,\xi_m) - q(x)\right].\end{aligned} \tag{20}$$

Again, we conjecture a linear solution to the differential equation. Solving, we obtain:

$$q^P(x,\xi_m) = \frac{x}{r+\delta-\mu}\left[\begin{array}{c} 1-\tau_P + \frac{\lambda_E\theta}{r+\delta+\lambda_E-\mu}\left[\tau_P - \overline{\tau} + \frac{\lambda_I[\overline{\tau}-\mathbb{E}(\tau_I|\xi_m)]}{r+\delta+\lambda_I-\mu}\right] \\ +\frac{\lambda_E(1-\theta)}{r+\delta+\lambda_E-\mu}\left[\tau_P - \underline{\tau} + \frac{\lambda_I[\underline{\tau}-\mathbb{E}(\tau_I|\xi_m)]}{r+\delta+\lambda_I-\mu}\right]\end{array}\right]. \tag{21}$$

Consider finally the Exogenous Policy Economy. The only necessary modification to the preceding analysis is that we must make the following substitution:

$$\mathbb{E}[\tau_I|\xi_m] \to \tau_I^{EX}. \tag{22}$$

In the Exogenous Policy Economy, the predetermined long-term tax rate τ_I^{EX} replaces industry-specific policy beliefs.

3. Inference in Economy-Wide Natural Experiments

This section considers causal parameter inference in the context of Economy-Wide Natural Experiments (EWNEs).

3.1. Two Numerical Examples

The econometric challenges in EWNEs are initially best illustrated by considering numerical examples. All the numerical examples assume: $M = 2$; $r = .05$; $\delta = .10$; $\mu = 0$; $\lambda_E = .15$; $\lambda_I = .15$; and $x = 1$. It is assumed $\overline{\tau} = .405$ and $\underline{\tau} = .21$. This parameterization is motivated by recent corporate taxation in the U.S. In 2017, President Trump signed into law what was viewed as a large cut in the corporate tax rate from 34% down to 21%. The chosen parameterization assumes the Trump tax reduction of 13 percentage points was the largest feasible, and that a tax increase one-half that size, or 6.5 percentage points, is the maximum feasible. It is assumed $\lambda^{-1} = 6.7$ years, approximating reality for major tax changes. Finally, the EWNE numerical examples assume $\nu = 2$, implying investment (equation (5)) is linear in q.

Consider first the following natural experiment. During the Pre-Experiment Stage $\tau_P = .34$, as was the case prior to the Trump tax cut. When the Experiment Stage arrives at an exogenous random date, the minimum feasible rate $\tau_E = .21$ will be applied to all firms ($\theta = 0$). In this context, we consider inference by an academic econometrician outside the government. Figure 2 depicts outcomes. The horizontal axis represents the causal parameter to be inferred, investment responsiveness ξ. The vertical axis measures the industry-specific investment increase when the experimental tax cut occurs.

It is useful to first contrast shock responses in the Endogenous versus Exogenous policy economies (thick solid schedule versus dashed schedule). The Endogenous Policy Economy follows the Baseline specification (see Section 2). Here the government will make the tax cut permanent during the Implementation Stage if average ξ exceeds $\xi^* = .50$. In the Exogenous Policy Economy, it is assumed government cannot reverse the tax cut. Notice, in this second economy, econometric inference serves no purpose.

As shown in Figure 2, firms will increase investment by more in response to the tax cut the higher their investment responsiveness parameter ξ. This is due to two factors. First, there is the direct effect, with equation (8) showing that for any change in q, the investment response increases linearly with ξ. Second, in the Endogenous Policy Economy, firms with higher ξ assess a higher probability of the tax cut being made permanent (see Lemma 1 and Figure 1), implying a larger increase in q.

The academic econometrician here can invert the observed response to infer the true causal parameter value, but only if she correctly accounts for policy feedback. If the econometrician is naïve, policy feedback is ignored and she instead uses the dashed line to perform inference. Notice, the dashed line represents a counter-factual economy in which the tax cut is permanent with probability 1. The naïve econometrician's parameter estimates are generally biased downward. For example, if $\xi = .40$, investment increases by .05. Incorrectly reading off the dashed line, the naïve econometrician will infer $\xi = .20$, a downward bias of 50%. Intuitively, in the Endogenous Policy Economy, agents understand the government will learn from the data, implying that the current policy may well be reversed, so the increase in q is small. The naïve econometrician incorrectly imputes the sluggish investment response to low values of ξ.

This situation has a close parallel in the investment literature. In particular, "implausibly low" estimates of ξ based upon *imputed* changes q have plagued empirical investment work going back to Summers (1981). A potential cause is that imputed q values are often based upon assuming firms view each tax change as permanent. For example, see Cummins, Hassett and Hubbard (1994). Such an inference procedure is internally inconsistent, since the primary purpose of the econometric analysis is to inform future endogenous tax policy decisions.

Correct causal parameter inference here also requires a correct stipulation of the government objective function. To illustrate, suppose ξ^* is raised from .50 to .75 following the High Cutoff scenario from Section 2. As shown in the dotted-dashed line in Figure 2, for each value of the causal parameter, the investment increase will be smaller under the higher cutoff. The failure to correctly account for the change in governmental objectives would result in downward bias. For example, if $\xi = .66$, the investment increase will be .054. Incorrectly reading off the original solid curve, the econometrician would infer $\xi = .44$, a downward bias of 33%. Intuitively, firms increase investment by a small amount because the true increase in q is small. Here the econometrician erroneously infers the smaller investment increase is due to a smaller ξ.

A frequent concern expressed by applied micro-econometricians is that selection limits external validity. For example, there might be concern that a tax cut is more likely to be tested by a government that knows it enjoys an unusually benign environment (high $\boldsymbol{\xi}$). By construction, we consider economies free from such selection/endogeneity problems (Assumption 1). Nevertheless, endogenous government decisions *post-experiment* will lead to problems of interpretation. To illustrate, consider an economy otherwise equivalent to the High Cutoff economy, with the exception that prior beliefs are less favorable to a tax cut post-experiment, with F being a triangular distribution on $[0, 1]$ with mode 0. Beliefs under this Negative Priors scenario are shown in Figure 1. Here firms assess a low probability of the tax cut being maintained long-term. Consequently, as shown in Figure 2, for any given value of the causal parameter ξ, the investment increase is smaller in this economy. Failure to account for this would lead to biased parameter inference.

The next example illustrates how ex post policy endogeneity creates potential for faulty inference of causal effect *signs*. Consider the same setting as in Figure 2, but with $\tau_P = .275$ rather than 34%. Thus, the experimental tax cut now being considered is from 27.5% down to 21%. This tax cut of 6.5 percentage points is half the size of that in the preceding example.

Casual intuition would suggest that the response to a tax cut one-half the size would lead to a halving of the investment response. However, Figure 3 shows that this is not the case. In particular, notice that in all the economies featuring optimal policy-setting post-experiment, the investment response to this small tax cut is actually negative for small values of ξ. In fact, in the Negative Priors economy, firms would always respond to the tax cut by decreasing their investment. Intuitively, the experimental tax cut is viewed as positive news only if firms believe there is a sufficiently high probability of the experimental evidence convincing the government to implement low taxes long-term.

Parameter inference based upon an EWNE requires the investment response be a monotone invertible function R, with:

$$\xi = R^{-1}\left[R(\xi)\right]. \tag{23}$$

Returning to Figure 3, recall we consider Rational Expectations Equilibria in which the government is assumed to be capable of inferring the causal parameters based on some set of econometric evidence at its disposal. However, it is apparent from Figure 3 that EWNE shock responses by themselves may be insufficient for this purpose. To see this, notice that in all three economies with endogenous policy-setting post-experiment, the investment response is non-monotone in ξ. Consequently, on some regions, the investment response outcome variable cannot be inverted to solve for the true causal parameter value. Here the econometrician would need to examine the components of the investment reaction (R), investment just before and just after the start of the experiment to distinguish alternative ξ values yielding the same R value.

3.2. Analytical Treatment of EWNEs

The change in investment at the start of the experiment is

$$R(\xi) = \left(\frac{\nu-1}{\nu}\right)^{\nu-1} \times \xi \times \left[\left(q^{E}(x,\xi)\right)^{\nu-1} - \left(q^{P}(x,\xi)\right)^{\nu-1}\right], \tag{24}$$

where q during the Experiment and Pre-Experiment stages are as shown in equations (18) and (21), respectively. When inspecting equation (24), it is important to note that beliefs (β) enter as arguments into the respective expressions for q. Any change in economic environment

(F or ξ^*) changing belief functions (Figure 1) changes the response function R. If $\nu = 2$, then:

$$R(\xi) = \frac{\xi}{2}\left[\frac{x}{r+\delta+\lambda_E-\mu}\right]\left[\begin{array}{c} \tau_P - \tau_E + \\ \frac{\lambda_I}{r+\delta+\lambda_I-\mu}\left[\tau_E - \beta(\xi)\underline{\tau} - (1-\beta(\xi))\overline{\tau}\right] \end{array}\right]. \tag{25}$$

The preceding equation shows industries will respond differently to a tax experiment due to differences in their ease of stock adjustment (ξ) as well as endogenous differences in beliefs about long-term tax policy ($\beta(\xi)$). It is also apparent that a correct stipulation of the belief function β is required to correctly infer the true causal parameter ξ based upon measured responses to experimental tax shocks. In turn, as shown in Section 2, correct stipulation of beliefs requires correct stipulation of prior beliefs (F) and correct stipulation of the government policy cutoff (ξ^*).

Computing differences in response functions across economies with endogenous versus exogenous policies, we find:

$$\underset{Endogenous}{R(\xi)} - \underset{Exogenous}{R(\xi)} = \frac{\xi}{2}\frac{\lambda_I\left[\tau_I^{EX} - \beta(\xi)\underline{\tau} - (1-\beta(\xi))\overline{\tau}\right]x}{(r+\delta+\lambda_E-\mu)(r+\delta+\lambda_I-\mu)}. \tag{26}$$

It follows from the preceding equation that there is necessarily a wedge between the experiment response functions across the Exogenous Policy and Endogenous Policy economies.

The following proposition summarizes the results of this subsection.

Proposition 2 *Response functions* (R) *for economy-wide tax shocks differ according to whether evidence is relevant (endogenous policy post-experiment) or irrelevant (exogenous policy post-experiment). Across economies with endogenous tax policies post-experiment, response functions vary with differences in government decision criteria* (ξ^*) *and differences in prior beliefs* (F) *regarding the distribution of causal parameters to be estimated.*

3.3. Observer Effects in EWNEs

This subsection considers the nature of feedback between the perceived credibility of evidence from an EWNE and the nature of the evidence itself.

Consider two economies facing the same experimental tax change, with the governments in the two economies both able to choose long-term policy. The two governments differ regarding the perceived credibility of experimental evidence. In Economy C, the government views experimental evidence as credible. In Economy NC, the government views experimental evidence as non-credible. Finally, suppose observation of firm behavior is either technologically feasible or not, and that all agents know whether or not observation is feasible when the model opens.

In Economy NC, long-term policy will be invariant to whether or not observation occurs. If observation occurs, the government in economy NC ignores the evidence and implements the policy that is optimal given prior beliefs, call it τ^*_{PB}. And if no observation occurs, the government has no choice but to rely on its priors, so it again implements τ^*_{PB}.

In Economy C, the probability distribution of the tax rate varies according to whether or not observation occurs. If no observation occurs, the government is forced to rely on priors, implementing $\tau_I = \tau^*_{PB}$. However, if observation occurs, the government views the experimental evidence as credible and uses it to infer $\boldsymbol{\xi}$. Firms then anticipate the government implementing $\tau^*_I(\boldsymbol{\xi})$ as defined in equation (12). Thus, they form industry-specific beliefs in accordance with β (equation (15)). An observer effect then arises from the change in the probability distribution of the tax rate resulting from the act of observation. Such observer effects for treated groups are known as *Hawthorne Effects.*[6]

The Hawthorne Effect for EWNEs can be expressed in terms of the change in q as follows:

$$\underset{\text{Observed}}{\left[q^E(x,\xi)-q^P(x,\xi)\right]} = \underset{\text{Not Observed}}{\left[q^E(x,\xi)-q^P(x,\xi)\right]} + \underset{\text{Hawthorne Effect}}{\frac{\lambda_I\left[\tau^*_{PB}-\beta(\xi)\underline{\tau}-(1-\beta(\xi))\overline{\tau}\right]x}{(r+\delta+\lambda_E-\mu)(r+\delta+\lambda_I-\mu)}}. \tag{27}$$

We have the following proposition.

Proposition 3 *If the government views the Economy-Wide Natural Experiment as credible (non-credible), the outcome variable, the change in investment during the experimental treatment period, is (not) contaminated by an observer effect.*

---

[6] The label is due to studies of lighting levels and worker productivity at Hawthorne Works.

As a numerical example, return to Figure 2, which illustrated responses to an experimental tax cut. Suppose now the econometrician lives in Economy C, with the Baseline assumptions operative, and where observation may or may not occur. In such an economy, the thick solid line in Figure 2 captures the investment response function if firms are observed. The dashed line in Figure 2 captures the response if firms are not observed. If firms are not observed, the government sets long-term policy based upon prior beliefs and maintains the tax cut with probability one, since the cutoff is just equal to the unconditional average of ξ. If firm actions are observed, sophisticated analysis allows the government to correctly infer $\boldsymbol{\xi}$ and so it implements a contingent optimal policy, with $\tau_I = \tau_I^*(\boldsymbol{\xi})$. Notice, if firms are observed, they increase investment by a small amount, especially if their industry-specific ξ is low, since in this case they attach a relatively low probability to the tax cut being continued in the long-term.

4. Controlled Natural Experiments

This section considers econometric inference in the context of Controlled Natural Experiments (CNEs).

4.1. A Numerical Example

The econometric challenges in CNEs are initially best illustrated by considering a specific numerical example. For this purpose, we retain the same parametric assumptions as in Section 3, but no longer restrict attention to quadratic adjustment costs, instead assuming $\nu = 8$, so that investment is not simply linear in q.

Consider the following CNE. During the Experiment Stage half the firms are randomly assigned to the minimum tax rate of 21% with the remaining half assigned to the maximum tax rate of 40.5%. The test statistic, denoted Δ, is the difference between mean investment for the control (low tax) and treatment (high tax) groups, controlling for industry:

$$\Delta(\xi) \equiv \underbrace{i^*[\underline{q}^E(x,\xi)]}_{Control} - \underbrace{i^*[\overline{q}^E(x,\xi)]}_{Treatment}. \tag{28}$$

Equation (18) pins down q during the Experiment Stage. Applying this equation, we have:

$$\begin{aligned} \underline{q}^E(x,\xi) &= \frac{x}{r+\delta-\mu}\left[1-\underline{\tau}+\frac{\lambda_I[\underline{\tau}-\mathbb{E}(\tau_I|\xi)]}{r+\delta+\lambda_I-\mu}\right] \\ \overline{q}^E(x,\xi) &= \frac{x}{r+\delta-\mu}\left[1-\overline{\tau}+\frac{\lambda_I[\overline{\tau}-\mathbb{E}(\tau_I|\xi)]}{r+\delta+\lambda_I-\mu}\right]. \end{aligned} \tag{29}$$

With expressions for q in-hand, investment (i^*) is again pinned down by equation (8).

Figure 4 plots the results. On the horizontal axis is the true value of the unknown causal parameter ξ, which measures investment responsiveness to changes in q. On the vertical axis is Δ. Since firms in the same industry will have identical optimal investment prior to the experiment, the vertical axis also measures the difference in differences. The thick solid line measures the control-treatment investment difference in the Endogenous Policy Economy which follows the Baseline parameterization for beliefs and the dashed line measures the difference in the Exogenous Policy Economy where, as in Section 3, it has been assumed that the tax rate will go to 21% for all firms after the experiment, with the government powerless to change this. The dotted dashed line considers the High Cutoff scenario ($\xi^* = .75$). The thin solid line considers the Negative Priors beliefs specification.

Figure 4 allows one to contrast the inference that will be made by a sophisticated econometrician, who accounts for the role of estimation in policy-setting, versus a naïve econometrician who ignores it. Consider, say, an academic econometrician working in the Endogenous Policy Economy where inference informs policy. If the econometrician is sophisticated, she will account for the link between policymaking and empirical evidence and use the thick solid line in performing inference, resulting in correct estimation of the parameter ξ. If the econometrician is naïve, she ignores the link and instead uses the dashed line to perform inference. Apparently, the naïve econometrician will understate the true value of the causal parameter

ξ. For example, suppose $\xi = .60$, resulting in an observed control-treatment investment difference of $11,500$. Incorrectly working along the dashed line, the naïve econometrician will infer this difference resulted from $\xi = .40$, a downward bias of 33%.

It is also apparent from Figure 4 that correct causal parameter inference in this CNE is contingent upon a correct stipulation of the government objective function. To see this, suppose the government opts for a higher threshold $\xi^* = .75$. Then the control-treatment investment difference changes from the thick solid line to the dotted-dashed line. If the econometrician failed to account for this change in objectives, bias would result. For example, suppose $\xi = 1$, resulting in an observed control-treatment investment difference of $17,500$ in the High Cutoff scenario. Incorrectly working along the thick solid line in Figure 4, the naïve econometrician will infer this difference resulted from $\xi = .79$, a downward bias of 21%.

An oft-mentioned concern for applied microeconometricians is that inferences regarding policy impacts will be non-representative if predicated upon a discretionary CNE. Assumption 1 rules out this type of selection bias. However, ex post endogeneity gives rise to a similar problem. To see this, suppose an academic econometrician examines the control-treatment investment difference in the Endogenous Policy Economy endowed with negative priors. Then the control-treatment investment difference changes from the thick solid line to the thin solid line. If the econometrician failed to account for this change in environment, biased inference would result. For example, suppose $\xi = 1$, resulting in an observed control-treatment investment difference of $14,000$. Incorrectly working along the thick solid line in Figure 4, the naïve econometrician will infer this difference resulted from $\xi = .67$, a downward bias of 33%.

4.2. Observer Effects in CNEs

This subsection considers the potential for control and treatment groups to exhibit observer effects in CNEs. To illustrate, we return to the same CNE and parameter values as in the preceding subsection, focusing on a government that is willing to use evidence from the CNE to set long-term tax rates. But observation may not be feasible. If observation is feasible, the government will set the long-term tax rate optimally based upon the econometric evidence. If observation is not feasible, the government will implement the minimum tax rate, which is optimal given prior beliefs.

The results are shown in Figure 5. On the horizontal axis is the true value of the causal parameter ξ. The plot contains the control-treatment group investment difference (solid lines), low tax group investment (dashed lines), and high tax group investment (dotted lines). In each case, the thicker line represents outcomes under observation and the thinner line represents outcomes if observation does not occur.

If observed, firms anticipate the government will utilize the experimental evidence in order to correctly infer $\boldsymbol{\xi}$, going on to implement the optimal tax rate $\tau_I^*(\boldsymbol{\xi})$. In contrast, absent observation, firms know the government will rely upon prior beliefs and will be certain that $\tau_I = \underline{\tau}$. Apparently, as shown in Figure 5, changes in the probability distribution of the tax rate post-experiment, resulting from observation, induce changes in investment by both treatment and control groups. In other words, here we have observer effects for both the treatment (Hawthorne Effect) and control groups (John Henry Effect).[7]

Finally, as shown in Figure 5, the act of observation changes the test statistic here, the control-treatment investment difference. The next subsection discusses why this is the case.

4.3. Analytical Treatment of CNEs

From equation (29) it follows that the difference between q across firms assigned low versus high tax rates is

$$\underline{q}^E(x,\xi) - \overline{q}^E(x,\xi) = \frac{(\overline{\tau} - \underline{\tau})x}{r + \delta + \lambda_I - \mu}. \tag{30}$$

---

[7] In labor studies, a John Henry Effect is said to arise if the control group exerts more effort to overcome their lack of treatment.

Notice, the preceding equation shows that the difference between q across control and treatment groups is actually invariant to the distribution of the policy variable post-experiment. Intuitively, random assignment ensures there is no selection based upon policy expectations. In fact, the difference in shadow values is simply the present value of the higher cash flow received by the low tax group during the experiment.

But recall, in our numerical example (Figure 4), the control-treatment investment difference actually changed with changes in expectations regarding the tax rate post-experiment. It is this investment difference that is the outcome variable observed by the econometrician, not the latent control-treatment shadow value difference (equation (30)). It is the behavior of investment, not shadow values, that the econometrician must understand.

From equations (8) and (30) it follows that the difference between the investment of low and high tax firms is:

$$\Delta(\xi) = \xi \left(\frac{\nu - 1}{\nu}\right)^{\nu-1} \left(\frac{x}{r + \delta - \mu}\right)^{\nu-1} \left[\begin{array}{c} \left(1 - \underline{\tau} + \frac{\lambda_I[\underline{\tau} - \mathbb{E}(\tau_I|\xi)]}{r+\delta+\lambda_I-\mu}\right)^{\nu-1} \\ -\left(1 - \overline{\tau} + \frac{\lambda_I[\overline{\tau} - \mathbb{E}(\tau_I|\xi)]}{r+\delta+\lambda_I-\mu}\right)^{\nu-1} \end{array} \right]. \tag{31}$$

A key point to note in equation (31) is that beliefs about tax policy post-experiment influence the investment of both the treatment and control groups, consistent with Figure 5. Since the act of observation influences beliefs regarding long-term policy, there will be observation effects for both the treatment group (Hawthorne Effect) and the control group (John Henry Effect). Moreover, the size of these effects will vary with prior beliefs and the parameters of the government objective function into which the evidence is fed.

Despite the existence of policy expectation effects for both treatment and control groups, it might be hoped that these effects will be of equal size across the two groups, so that the control-treatment investment difference will be left uncontaminated. However, as shown in equation (31), expectation effects do not generally cancel. In fact, it is instructive to consider the exception proving the rule. If $\nu = 2$, investment is linear in q and the observation effects cancel. In particular, it follows from equation (31) that:

$$\nu = 2 \Longrightarrow \Delta(\xi) = \frac{\xi}{2} \frac{(\overline{\tau} - \underline{\tau})x}{r + \delta + \lambda_I - \mu} \tag{32}$$

If $\nu = 2$, then Δ is invariant to expectations regarding the distribution of the policy variable post-experiment. Thus, in the special case of quadratic investment costs, the control-treatment investment difference is uncontaminated by ex post endogeneity.

In order to provide a more complete characterization of the circumstances under which differences (and differences in differences) derived from CNEs are immune from post-experiment policy expectations contamination, consider the following General Investment Cost Function.

Definition 4 *General Investment Cost Function: The fixed cost to positive investment is* $\varphi^+ \geq 0$. *The fixed cost to negative investment is* $\varphi^- \geq 0$. *Capital can be purchased at price* P^+ *and sold at price* $P^- \leq P^+$. *Adjustment costs are* ψ, *where* ψ *is a strictly convex twice differentiable function of investment attaining a minimum value of zero at* $i = 0$.

Since the General Investment Cost Function shares with the initially-posited cost function (equation (3)) the property of being invariant to k, it follows that the shadow value formulae derived above (Subsection 2.4) remain valid. Under such a cost function, investment is weakly monotone increasing in q. Further, absent fixed costs, investment is continuous in q, with $i^* = 0$ optimal for all $q \in [P^-, P^+]$, turning negative at points to the left of this interval and positive at points to the right. With fixed costs, optimal accumulation is zero over a wider interval of q values, and exhibits discontinuities at the optimal thresholds for switching from inaction to action.

Recall, under the initially-posited investment cost function, the control-treatment investment difference (as well as difference in differences) was shown to be invariant to post-experiment policy variable expectations if and only if investment is linear in q, which held if $\nu = 2$. To ensure that investment is linear in q under a General Investment Cost Function, one must rule out fixed costs, wedges between the buy and sell price of capital, and assume quadratic adjustment costs. We thus have the following proposition.

Proposition 5 *If and only if the CNE is relevant (endogenous policy post-trial), both the treatment and control group will exhibit observer effects. The difference between control and treatment group investment (and the difference in their differences) is invariant to factors affecting post-experiment policy variable expectations if and only if the General Investment Cost Function features: a quadratic adjustment cost function* (ψ)*; zero fixed costs* ($\varphi^{-} = \varphi^{+} = 0$)*; and zero wedge between the buy and sell price of capital* ($P^{-} = P^{+}$).

The preceding proposition makes clear that even ideal controlled natural experiments must be understood as being predicated upon strong functional form assumptions. To illustrate this, it is worth noting that a pre-requisite for correct parameter inference based upon an observed control-treatment difference is that the difference function Δ be strictly monotone in ξ. Formally, parameter inference based on an CNE entails inverting the difference, allowing one to write:

$$\xi = \Delta^{-1} \left[\Delta(\xi) \right]. \tag{33}$$

With this in mind, suppose investment costs are quadratic ($\nu = 2$). If there are no fixed costs and the buy and sell price of capital are equal, then equation (32) still applies, with the difference between control and treatment group investment being linear in ξ. This scenario is depicted in the dotted line in Figure 6, which considers our original CNE but now assumes a capital goods price equal to 5. Here the test statistic can be inverted to solve for the causal parameter ξ.

The solid line in Figure 6 assumes the sell price of capital is 4.25, a bit lower than the buy price of 5. As shown, this real friction would create major problems in terms of causal parameter inference. In particular, for ξ on the interval between .40 and .60, the investment difference is zero, with both the low and high tax firms finding zero investment to be optimal on respective regions of optimal inaction. The result is that the causal parameter is not identified by the CNE test statistic. Here the econometrician would need to look at data other than investment, such as average q values, to distinguish alternative ξ values.

5. Conclusion

This paper illustrates an inherent tension between the credibility of empirical estimates derived from natural experiments and their practical utilization. In particular, once this econometric methodology is perceived as being sufficiently credible, estimates derived from it will be used in setting policy. But this contaminates the original econometric estimates by exposing them to ex post endogeneity, with treatment responses dependent upon: the policymaker objective function into which estimates are fed; prior beliefs regarding the causal parameters to be estimated; and endogenously heterogeneous policy expectations. As shown, the failure to account for ex post endogeneity leads to faulty inference regarding the signs and magnitudes of causal parameters. Far from being stand-alone objects, correct causal interpretation of natural experiments may require an extremely subtle analysis and may require the imposition of strong functional form assumptions.

The quantitative examples illustrated that causal parameter estimates derived from natural experiments can suffer from large percentage biases if ex post endogeneity is not taken into account. Worse still, in some instances ex post endogeneity can lead to complete sign reversals and/or a complete lack of identification.

Of course, it might be objected that our rational expectations framework imputes too much sophistication to agents. However, even if one were to admit a departure from rationality, it is not clear whether this would decrease or increase bias. For example, subjective non-Bayesian updating can feature overreaction to signals in addition to underreaction to signals. A related objection that experimentalists may raise is that agents are simply not sophisticated enough understand the link between experiments and the policy setting process. To counter, we would argue that, whatever its present-day merits, this defense will become increasingly tenuous the closer experimentalists come to achieving their ultimate goal of systematic evidence-based policymaking. That is, any empirical methodology that relies upon systematically fooling agents cannot be credible in the long run.

The more general point illustrated by this paper is that how one interprets econometric estimates depends on whether and how those same estimates will be used subsequently. That is, there is an important difference between passive econometric estimation versus econometric exercises that are properly understood as joint estimation and control.

References

Abel, A., Eberly, J., 1997. An exact solution for the investment and market value of a firm facing uncertainty, adjustment costs, and irreversibility. Journal of Economic Dynamics and Control 21, 831-852.

Acemoglu, D., 2010. Theory, general equilibrium, and political economy in development economics. Journal of Economic Perspectives 24, 17-32.

Alti, A., 2003. How sensitive is investment to cash flow when financing is frictionless? The Journal of Finance 58, 707-722.

Angrist, J. D., Pischke, J.-S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Angrist, J. D., Pischke, J.-S., 2010. The credibility revolution in economics: How better research design is taking the con out of econometrics. Journal of Economic Perspectives 24, 3-30.

Chassang, S., Padro i Miguel, G., Snowberg, E., 2012. Selective trials: A principal-agent approach to randomized controlled experiments. American Economic Review 102, 1279-1309.

Cummins, J. G., Hassett, K. A., Hubbard, R. G., 1994. A reconsideration of investment behavior using tax reforms as natural experiments. Brookings Papers on Economic Activity 1994, 1-74.

Deaton, A., 2010. Instruments, randomization, and learning about development. Journal of Economic Literature 48, 424-455.

Gomes, J., 2001. Financing investment. American Economic Review 91, 1263-1285.

Greenstone, M., 2009. Toward a culture of persistent regulatory experimentation and evaluation. In D. Moss and J. Cisternino, eds. New Perspectives on Regulation, Cambridge, MA., The Tobin Project.

Heckman, J. J., 2000. Causal parameters and policy analysis in economics: A twentieth century retrospective. Quarterly Journal of Economics 115, 45-97.

Kocherlakota, N., 2018. Practical Policy Evaluation. NBER working paper number 24643.

Linde, J., 2001, Testing for the Lucas critique: A quantitative investigation, American Economic Review 91, 986-1005.

Ljungqvist, L., 2008. Lucas critique. The New Palgrave Dictionary of Economics, Second Edition. Edited by Steven N. Durlauf and Lawrence E. Blume.

Lucas, R. E., Jr., 1976. Econometric policy evaluation: A critique. in K. Brunner and A. Meltzer Eds., The Phillips Curve and Labor Markets, Amsterdam: North Holland.

Moyen N., 2005. Investment-cash flow sensitivities: constrained vs unconstrained firms. Journal of Finance 59, 2061-2092.

Romer, C., Romer, D., 2014. The NBER Monetary Economics Program. NBER Reporter.

Sargent, T. J., 1971. A note on the accelerationist controversy. Journal of Money Credit and Banking 3, 721-725.

Sargent, T. J., 1973. Rational expectations and the dynamics of hyperinflation. International Economic Review 14, 328-350.

Sargent, T. J., 1977. The demand for money during hyperinflations under rational expectations, I. International Economic Review 18, 59-82.

Sargent, T. J., 1984. Autoregressions, expectations and advice. American Economic Review 74, 408-415.

Sims C., 1982. Policy analysis with econometric models. Brookings Papers on Economic Activity, 107-64.

Summers, L., 1981. Taxation and corporate investment: A q-theory approach. Brookings Papers on Economic Activity 1, 65-105.

Taylor, J. B., 1979. Estimation and control of a macroeconomic model with rational expectations. Econometrica 47, 1267-1286.

Wolpin, K., 2013. *The limits of inference without theory.* MIT Press.

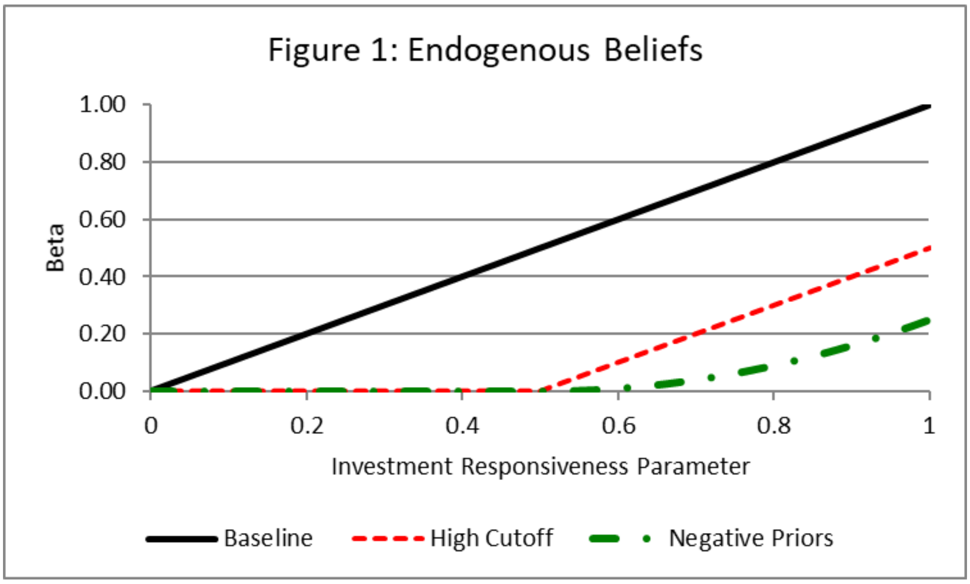


Figure 1: Endogenous Beliefs

This figure depicts the endogenous belief a firm will assign to the government choosing a low tax rate in the long-term, denoted as Beta in the model. As shown in the figure, the Beta variable will vary with the firm's Investment Responsiveness Parameter. The figure depicts three scenarios: Basline, High Cutoff and Negative Priors.

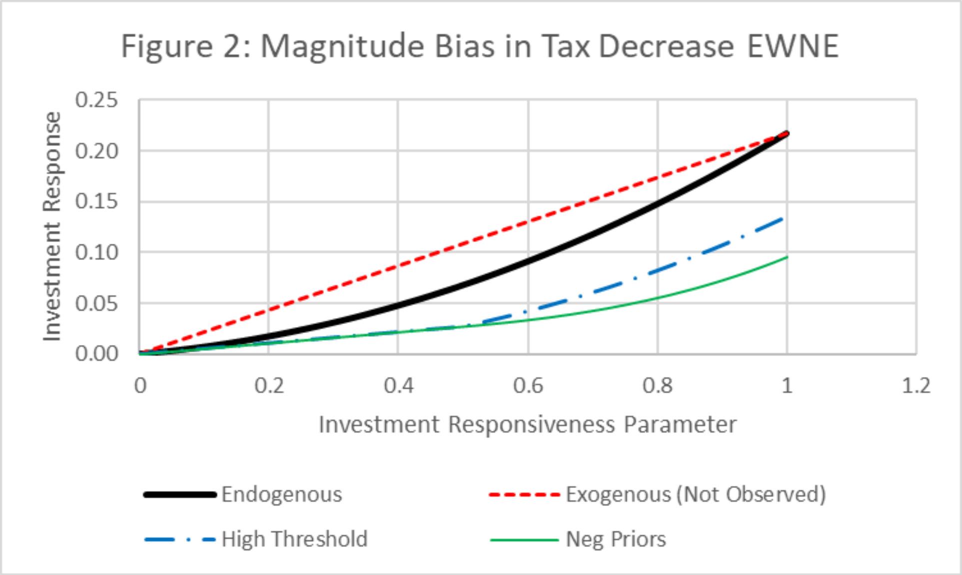
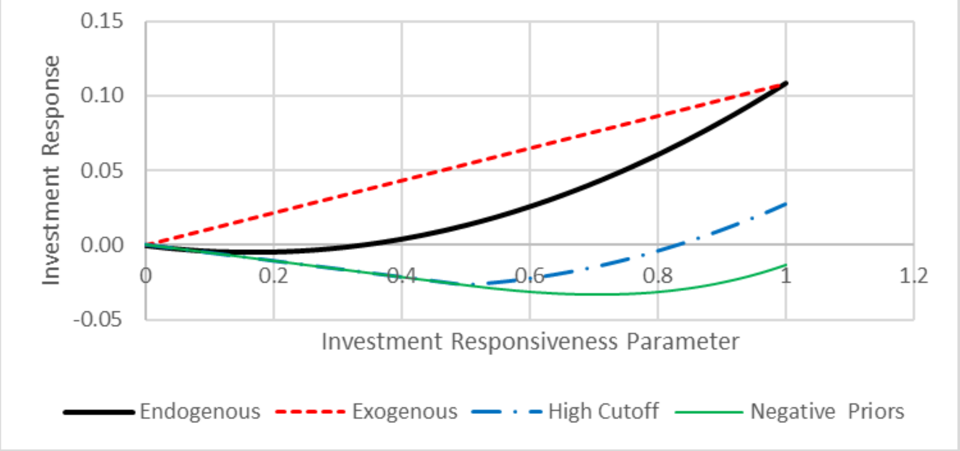


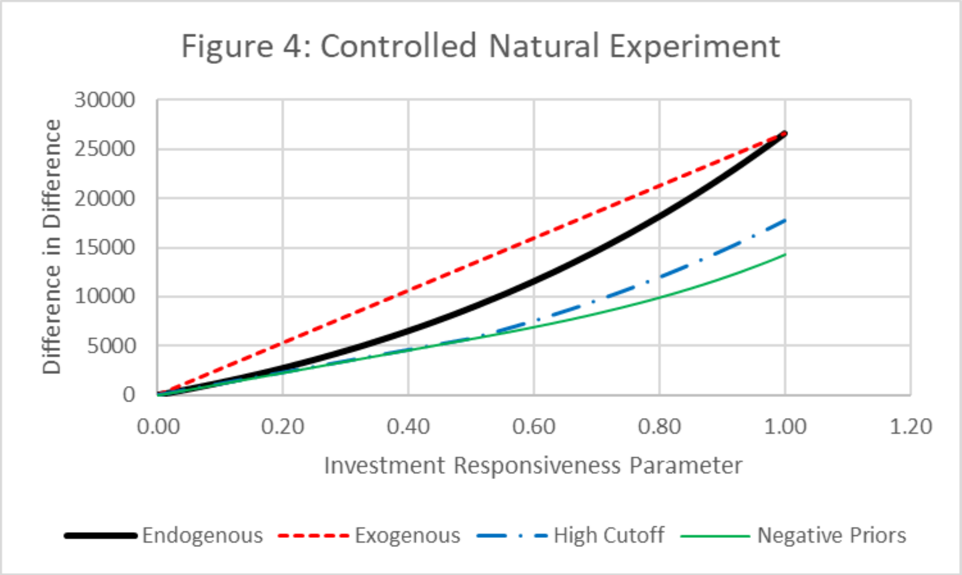
Figure 2: Magnitude Bias in Tax Decrease EWNE

This figure depicts the change in investment that will occur at the onset of a decrease in the tax rate from 34% to 21% under an Economy-Wide Natural Experiment (EWNE).

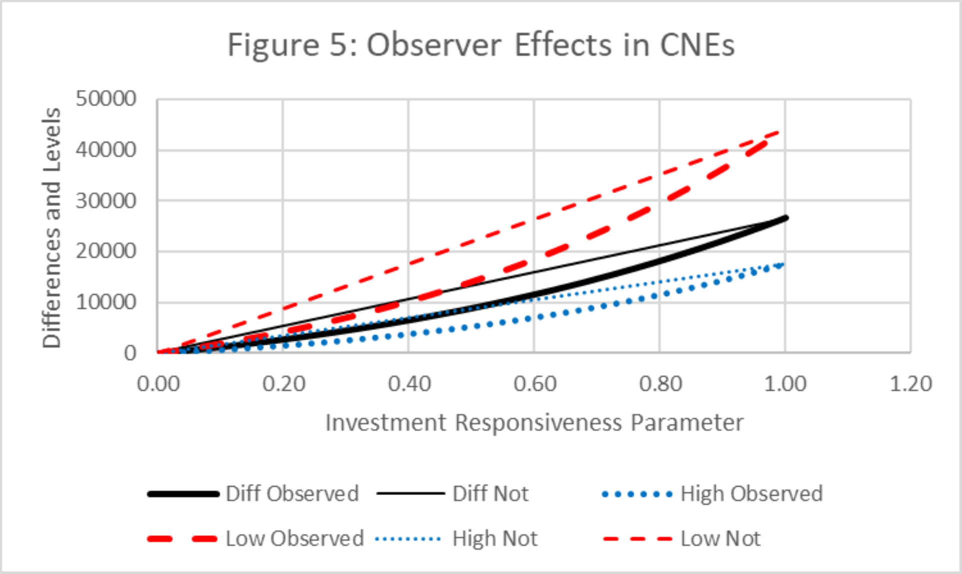
Figure 3: Sign Reversal and Non-Monotonicity in Tax Decrease EWNE



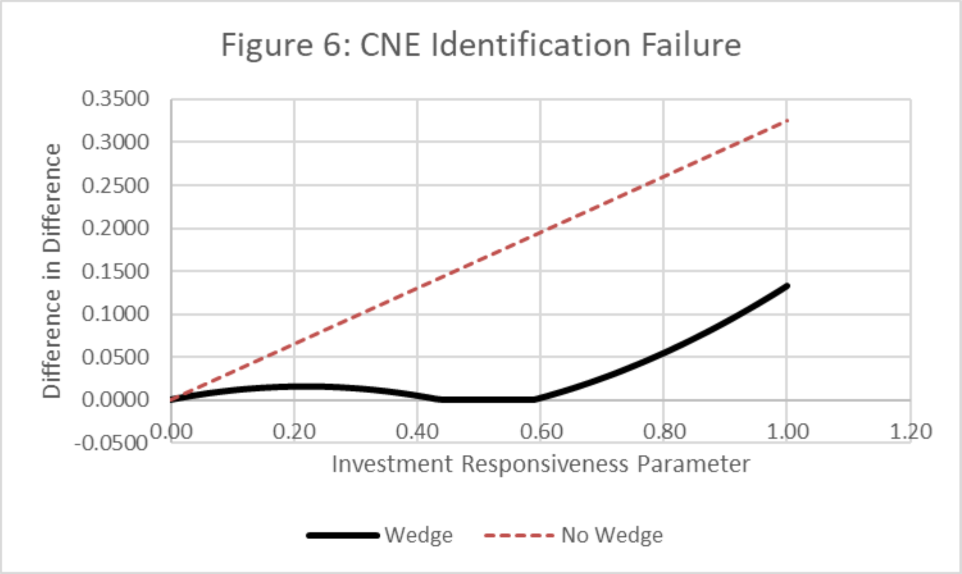
This figure depicts the change in investment that will occur at the onset of a decrease in the tax rate from 27.5% to 21% under an Economy-Wide Natural Experiment (EWNE).



This figure depicts the Difference in Difference in investment across firms treated with tax rates of 21% versus tax rates of 40.5%.



This figure depicts the level of investment and the difference in investment across high and low tax rate firms depending on whether the firms are observed or not. In addition



This figure depicts the Difference in Difference in investment across firms treated with tax rates of 21% versus tax rates of 40.5%. The Wedge Scenario considers that there is a wedge between the price at which capital is purchased versus sold.