

## EQUILIBRIUM COUNTERFACTUALS\*

BY GILLES CHEMLA  AND CHRISTOPHER HENNESSY*Imperial College, U.K.; DRM/CNRS; LBS, U.K.; CEPR; ECGI*

We incorporate structural modelers into the economy they model. Using traditional moment matching, they treat policy changes as zero probability (or exogenous) “counterfactuals.” Bias occurs since real-world agents understand policy changes are positive probability events guided by modelers. Downward, upward, or sign bias occurs. Bias is illustrated by calibrating the Leland model to the 2017 tax cut. The traditional identifying assumption, constant moment partial derivative sign, is incorrect with policy optimization. The correct assumption is constant moment total derivative sign accounting for estimation-policy feedback. Model agent expectations can be updated iteratively until policy advice converges to agent expectations, with bias vanishing.

## 1. INTRODUCTION

The strength of structural methods relative to reduced-form methods is their (potential) ability to overcome the rational expectations critique of Lucas (1976), who described pitfalls in extrapolating econometric estimates across policy environments. After all, the objective of structural inference is to estimate the magnitudes of policy-invariant parameters. With unbiased estimates of policy-invariant parameters in hand, alternative policies can be rigorously evaluated.<sup>1</sup>

The starting observation for this article is that, as argued immediately below, an important class of structural methods, those involving moment matching, violate rational expectations when the structural models serve their intended function of informing policy decisions with positive probability. With this observation in mind, *the primary objectives of this article are threefold: characterize the nature of bias that results from the violation of rational expectations when there is policy feedback in moment matching estimation; offer a feasible algorithm for eliminating the bias; and characterize how moment selection criteria change under the proposed algorithm.* Aside from early philosophical commentary from Sargent–Sims–Wallace (cited below) on potential internal inconsistencies in doing normative work under the rational expectations assumption, the structural moment matching literature has not yet addressed in an operational way any of these three challenges.

In order to see the violation of rational expectations, consider that in specifying the decision problem of agents inside her model, the structural econometrician must specify government policy. Critically, it is customary to parameterize models according to status quo policy (or to specify policy as an inviolable exogenous process). This standard practice violates rational expectations since the agents inside the model are treated as being ignorant of future

\*Manuscript received January 2020; revised August 2020.

We thank seminar participants at the CEPR ESSET Meetings, Minnesota, San Diego, LBS, Dallas, Georgia State, Imperial, Oxford, Dauphine, Frankfurt, Goethe, Lugano, IDC, Tel Aviv University, and Hebrew University. We also thank Lorian Mancini, Antonio Mele, and Antoinette Schoar for early feedback.

The copyright line for this article was changed on June 11, 2021 after original online publication. Please address correspondence to: Gilles Chemla, Imperial College Business School, London SW7 2AZ, U.K. E-mail: [g.chemla@imperial.ac.uk](mailto:g.chemla@imperial.ac.uk).

<sup>1</sup> See Sargent (2018) for a related discussion.

endogenous policy changes despite the goal of estimation being to inform endogenous policy decisions.

What are the actual implications of such violations of rational expectations (*nature of the bias*)? And, what can be done about them operationally (*moment selection criteria and algorithmic remedy*)? In order to address these three questions, we consider an economy in which “real-world” agents with rational expectations are placed alongside a structural econometrician who will give policy advice. That is, we develop a model that incorporates the structural modeler within it, consistent with the “communism of models” of Sargent (2010). The real-world agents are privately endowed with a policy-invariant parameter.<sup>2</sup> Knowledge of this parameter would be sufficient for the government to set policy at first-best. The econometrician will observe an empirical moment derived from agent actions, which will serve as the basis for her parameter inference.

The baseline model ingredients are as follows. When the model opens, government policy is set at a predetermined “status quo” value, denoted as  $\gamma_0$ .<sup>3</sup> Nature then draws the unknown parameter  $u$  from the real line.<sup>4</sup> There is a continuum of rational real-world agents who privately observe the parameter and choose their actions noncooperatively. The econometrician then observes a moment  $m$  that satisfies the *standard moment monotonicity condition*: The partial derivative  $\partial m/\partial u$  has a constant sign.<sup>5</sup> The econometrician then matches the model-implied moment with the real-world moment in order to draw an inference  $\hat{u}$ .<sup>6</sup> Finally, with positive probability the government will subsequently enjoy discretion to move the policy variable away from  $\gamma_0$  based upon the econometrician’s report of  $\hat{u}$ . A monotonic function  $g$  maps  $\hat{u}$  into government policy.

We begin by offering a complete characterization of the nature of biases that arise when real-world agents have rational expectations whereas the structural model violates rational expectations by treating changes in policy away from  $\gamma_0$  as zero probability (“counterfactual”) events.<sup>7</sup> As shown, given their knowledge of the true parameter  $u$ , rational expectations implies the real-world agents can correctly anticipate the econometrician’s inference  $\hat{u}(u)$ , and are able to anticipate errors driving a wedge between  $\hat{u}(u)$  and the true parameter  $u$ . More generally, rational expectation agents correctly anticipate the function  $\hat{u}(\cdot)$  determining the econometrician’s inference for each possible realized value  $u$ . Agents then correctly anticipate the policy the government will implement once it enjoys policy discretion, specifically  $g[\hat{u}(u)]$ .

As shown, the failure to impose rational expectations results in bias,  $\hat{u}(u) \neq u$ , at all points except the one possible realization of the unknown parameter, call it  $u_0$ , that would justify the government maintaining the status quo  $\gamma_0$ . Intuitively, in this exceptional case, the econometrician will not violate rational expectations since parameterizing the structural model as if the status quo will be implemented even when the government enjoys policy discretion is actually correct here.

The nature of bias depends upon the properties of the moment function  $m$  and the government policy function  $g$  mapping reports  $\hat{u}$  to discretionary policy. The central insight is as follows: If real-world agents have rational expectations, the empirical moment first varies directly with changes in the parameter  $u$ . This *direct effect*,  $m_u$ , is accounted for by the structural model. However, the real-world moment also varies indirectly due to the rational expectations of agents of feedback from the parameter inference ( $\hat{u}$ ) to discretionary government policy ( $g(\hat{u})$ ). Thus, the real-world empirical moment can be expressed as  $m[u, g(\hat{u}(u))]$ . It is the *indirect effect* arising from joint estimation and policy control,  $m_g g' \hat{u}'$ , that has traditionally been ignored in moment matching.

<sup>2</sup> Idiosyncratic parameters can be, say, multiples of a common aggregate parameter.

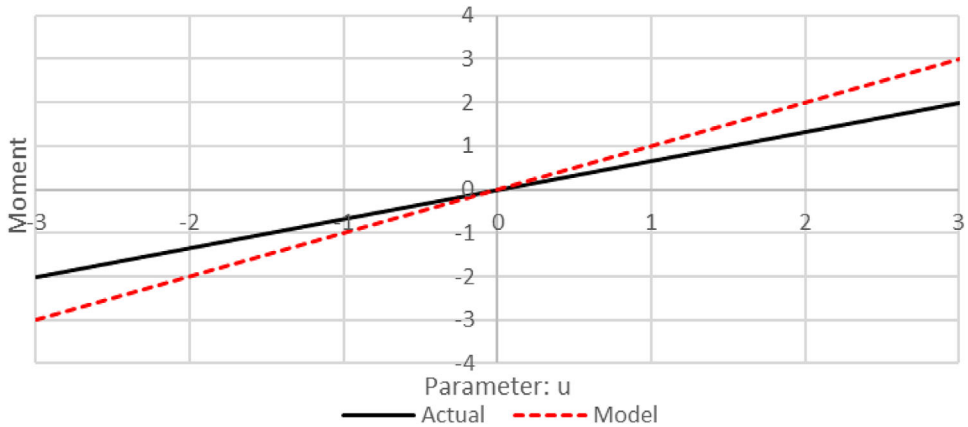
<sup>3</sup> As discussed below,  $\gamma_0$  can instead be policy in some state under exogenous Markovian policy.

<sup>4</sup> Vectors of unknowns are considered as an extension.

<sup>5</sup> See Gallant and Tauchen (1996) and Adda and Cooper (2003).

<sup>6</sup> This is in the spirit of Roberds (1987).

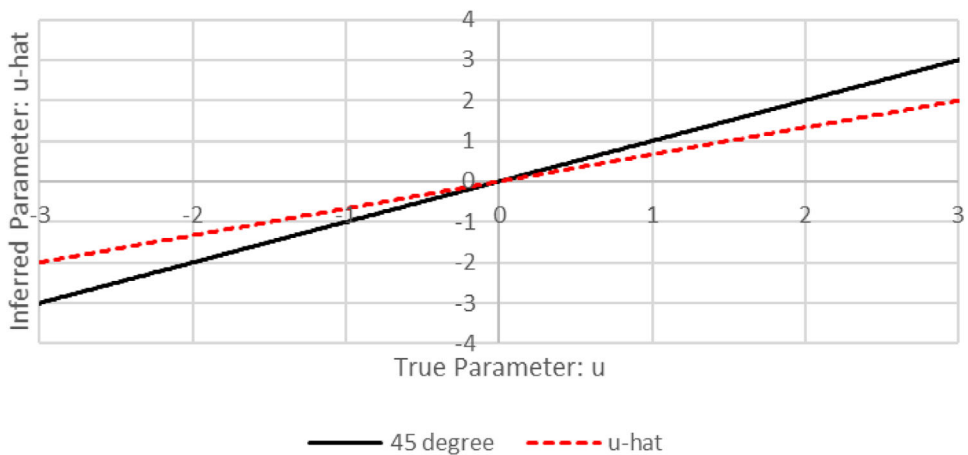
<sup>7</sup> Our argument also applies if  $\gamma_0$  is an exogenous element of a Markovian policy vector.



NOTE: Case:  $m(+, -), g(+)$ .

FIGURE 1

ACTUAL MOMENT VERSUS MODEL MOMENT  
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

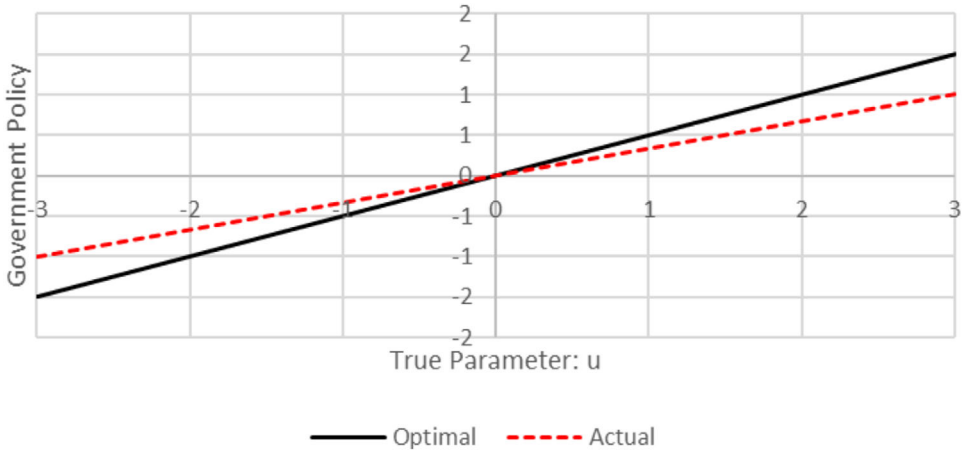


NOTE: Case:  $m(+, -), g(+)$ .

FIGURE 2

PARAMETER INFERENCE  
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

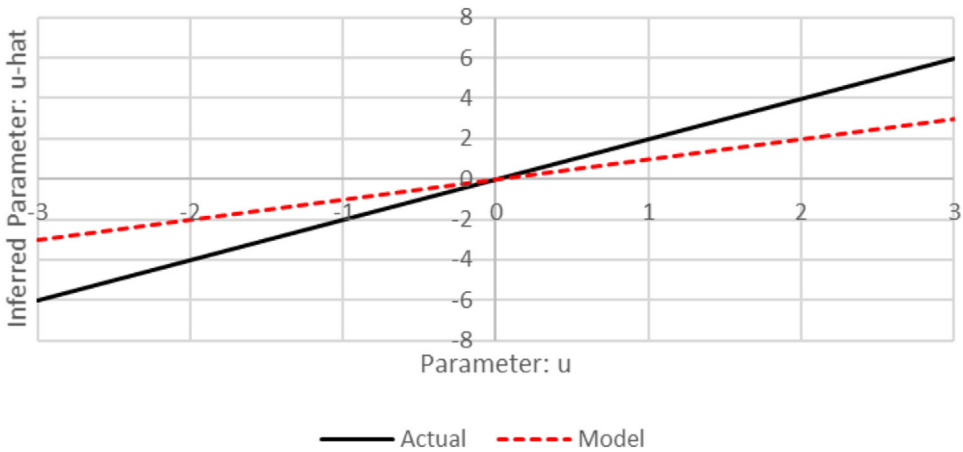
There are three cases to consider. In the first case, the sign of the indirect effect is opposite to that of the direct effect (see Figure 1). Here, as shown in Figure 2, the estimated parameter overshoots for  $u < u_0$  and then undershoots for  $u > u_0$  (and recall,  $u_0$  justifies the status quo policy  $\gamma_0$ ), resulting in policy bias (Figure 3). Intuitively, the modeler incorrectly treats small observed changes in the empirical moment to small changes in  $u$  because she here fails to account for the *countervailing* indirect effect. In the second case, the indirect effect is small in absolute value and has the same sign as the direct effect (see Figure 4). Here, as shown in Figure 5, the estimated parameter undershoots for  $u < u_0$  and overshoots for  $u > u_0$ , again resulting in policy bias (Figure 6). Intuitively, the modeler incorrectly treats large observed changes in the empirical moment to large changes in  $u$  because she here fails to account for the *amplifying* indirect effect. In the third case, the indirect effect is large in absolute value and has the same sign as the direct effect (see Figure 7). Here, as shown in Figure 8, the estimated parameter actually *decreases* with the true parameter, and it is possible that the



NOTE: Case:  $m(+, -), g(+)$ .

FIGURE 3

ACTUAL VERSUS OPTIMAL POLICY  
 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]



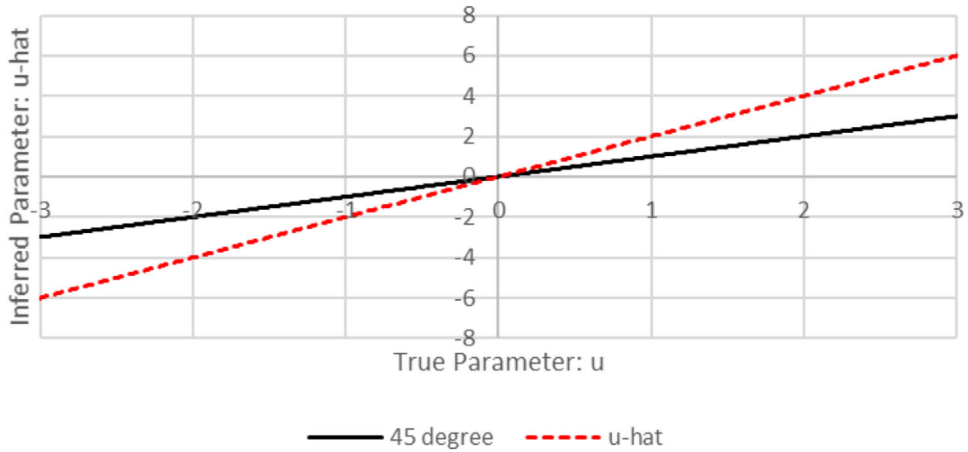
NOTE: Case:  $m(+, +), g(+)$ .

FIGURE 4

ACTUAL MOMENT VERSUS MODEL MOMENT  
 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

estimated parameter always has the wrong sign, for example,  $\hat{u}(u) = -u$ . The subtle intuition for this case is provided in the body of the article.

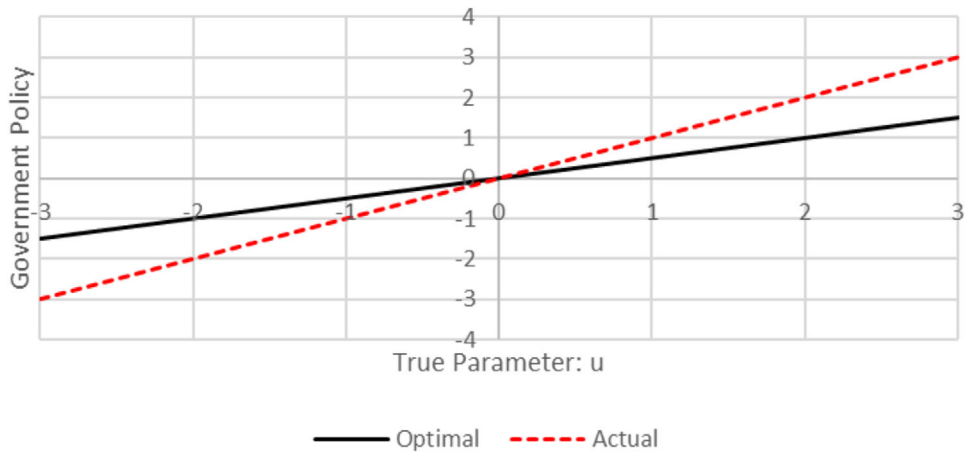
We illustrate the potential quantitative significance of these effects by considering an econometrician whose objective is to infer bankruptcy costs using the canonical structural model of Leland (1994). In particular, we consider the recent cut in the corporate income tax rate implemented under the Tax Cuts and Jobs Act of 2017. Here, the structural econometrician backs out implied bankruptcy costs from observed values of corporate interest coverage ratios. By assumption, the econometrician knows the underlying real technology but fails to impose rational expectations on firms inside the model. In our calibrated example, this leads to an eightfold overstatement of bankruptcy costs. Intuitively, firms rationally anticipate a tax cut and thus choose low leverage in light of the low value of future debt tax shields. Neglecting this fact, the econometrician mistakenly infers low leverage stems from extremely high bankruptcy costs.



NOTE: Case:  $m(+, +), g(+)$ .

FIGURE 5

PARAMETER INFERENCE  
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]



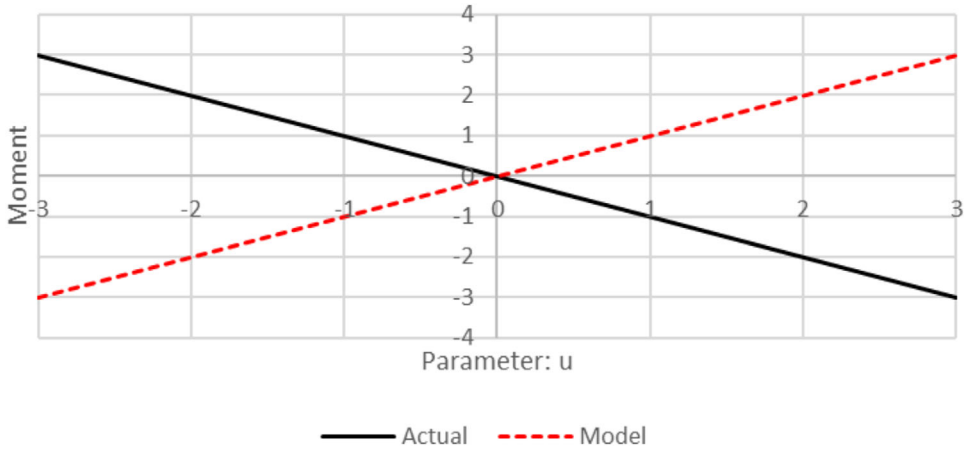
NOTE: Case:  $m(+, +), g(+)$ .

FIGURE 6

ACTUAL VERSUS OPTIMAL POLICY  
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

Importantly, we show that the nature of inference must change when the government will (with positive probability) engage in endogenous policy changes informed by estimation. That is, the econometric procedure must change as one moves from passive to active estimation. For example, with policy feedback, the standard moment monotonicity condition, which focuses on *partial derivatives* of moments, is neither necessary nor sufficient for correct structural parameter identification. Rather, we show that *total derivatives cum policy feedback* are the correct moment selection criterion in the context of joint estimation and control exercises. This implies that a moment that is informative (uninformative) under passive estimation may be uninformative (informative) under active estimation. That is, moment selection should vary according to whether estimation is active or passive.

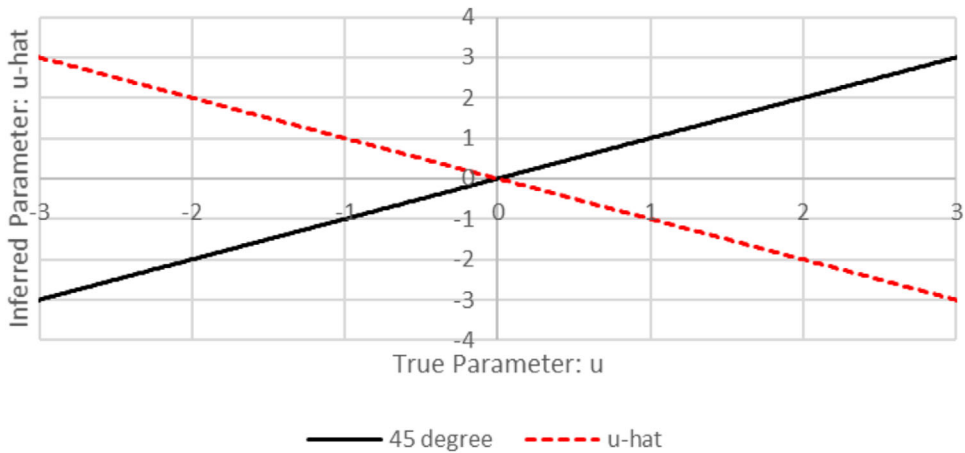
Based on the preceding insights, we develop a simple algorithmic procedure for achieving unbiased parameter estimates and first-best government policy. Recall, the underlying source



NOTE: Case:  $m(+, +), g(+)$ .

FIGURE 7

ACTUAL MOMENT VERSUS MODEL MOMENT  
 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

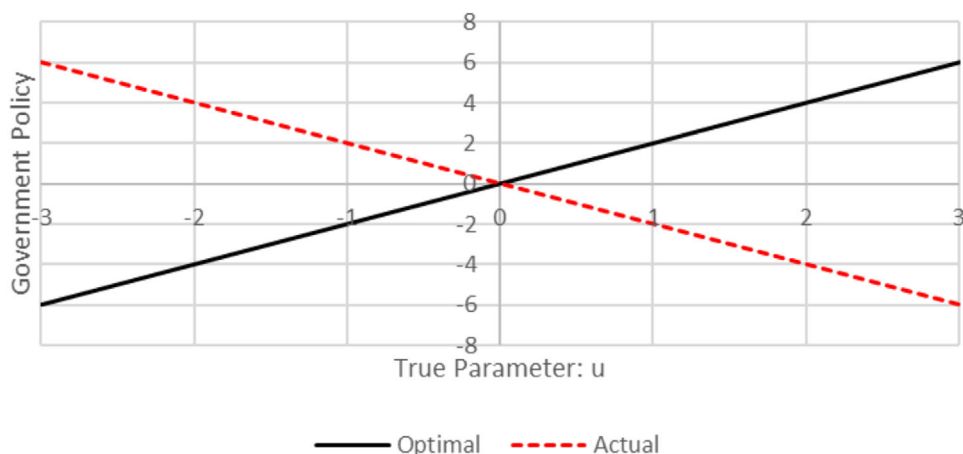


NOTE: Case:  $m(+, +), g(+)$ .

FIGURE 8

PARAMETER INFERENCE  
 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

of bias was that the agents inside the structural model were treated as being ignorant of possible endogenous policy changes whereas real-world agents have rational expectations and understand that endogenous policy changes are positive probability events. Thus, there was a systematic gap between model agent beliefs and real-world agent beliefs, a gap left open using standard moment matching procedures. However, this gap can be closed by iterating on inference and policy advice. The econometrician starts iteration  $n$  with a provisional policy recommendation  $\gamma_n$ . A corresponding parameter inference  $\hat{u}_n$  is derived by matching the observed real-world moment with the  $n$ th iteration model-implied moment  $m(\hat{u}_n, \gamma_n)$ . That is, in iteration  $n$  agents are treated as anticipating implementation of the provisional policy recommendation. Next, the implied optimal government policy  $g(\hat{u}_n)$  is computed and treated as the next provisional policy recommendation  $\gamma_{n+1}$ . Iteration proceeds until policy



NOTE: Case:  $m(+, +)$ ,  $g(+)$ .

FIGURE 9

ACTUAL VERSUS OPTIMAL POLICY  
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

convergence/internal consistency. That is, a fixed point is found where the policy advice supports the parameter inference, and vice versa.

Our article is in the spirit of Sargent and Wallace (1976), which contains an early discussion of the philosophical conundrums posed by attempting to impose rational expectations in the context of normative work. Sims (1982) makes the point more forcefully that there is an inconsistency in rational expectations econometric methods (specifically vector autoregressions [VARs]) that treat time series as reflecting optimal behavior on the part of private sector agents, but not the government. Sargent (1984) acknowledges these problems, as does Sargent (1987), stating, “There is a logical difficulty in using a rational expectations model to give advice, stemming from the self-referential aspect of the model that threatens to absorb the economic adviser into the model... That simultaneity is the source of the logical difficulties in using rational expectations models to give advice about government policy.” Our article differs from these papers in four respects. First, we focus specifically on moment matching methods. Second, we move beyond noting the existence of philosophical contradictions and instead characterize the nature of bias induced. Third, we propose a feasible corrective algorithm in the context of moment matching methods. Finally, we discuss how moment selection criteria change.

Building on the earlier work of Sims and Zha (2006), and Farmer et al. (2009), Bianchi and Ilut (2017) show how such internal contradictions can be avoided provided one confines attention to a specific *passive* form of counterfactual analysis. In this approach, historical and prospective government policy is modeled as an *exogenous* Markov chain whose probability law *cannot be changed*. Counterfactual analysis is then performed on particular historical periods by holding fixed nonpolicy shocks and then pretending as if the realization of the Markov chain differed from actual policy during the period of interest, for example, a counterfactual shift from passive to active monetary policy during the '60s and '70s. Strictly speaking, such an approach precludes the econometrician giving advice that can *actually alter the policy process*, as well as analysis of *novel policies*. After all, to escape the rational expectations trap here, the policy Markov chain must be treated as unalterable, otherwise the feedback bias we describe would emerge.

Also related to the present article is work by Chemla and Hennessy (2019), showing that a bias arises when quasi-experimental evidence is used to inform endogenous policy decisions. Arguably, the present article's critique is more problematic in that it is internal, taking

models and agent rationality seriously, a goal shared by many structural econometricians. Another important difference is that within the logic of a structural model, bias characterization is simpler. Finally, this article offers a feasible algorithm for avoiding bias and achieving first-best policy, again within the logic of a structural model. Despite these differences, the two papers share the message that the econometric tool kit changes as one moves from passive estimation, that influences policy with probability zero, to active estimation, that influences policy with positive probability.

Structural methods have been used across a wide variety of applied fields. In their influential paper, Kydland and Prescott (1996) advocate the use of calibrated structural models to evaluate policy alternatives, with subsequent macroeconomics and asset pricing papers treating moment matching more or less formally. For examples of structural methods in finance and banking, see Gomes (2001), Moyen (2005), and Hennessy and Whited (2005). Keane and Wolpin (2002) develop a granular structural model of public assistance programs. Adda and Cooper (2003) provide numerous applications of structural methods including labor and capital demand. Keane and Wolpin (1997), Rust (2013), and Wolpin (2013) provide overviews of structural applications in labor economics, public finance, and industrial organization, among others. Importantly, existing treatments are silent on how moment matching can proceed in a manner consistent with rational anticipation of policy advice.

The rest of the article is as follows. Section 2 describes the economic setting. Section 3 characterizes the nature of bias if the econometrician fails to impose rational expectations on agents inside the model. Section 4 shows how, under technical conditions, unbiased parameter inference and first-best government policy can be achieved through an iterative algorithm. In addition, Section 4 shows how traditional moment selection criteria are altered when one moves away from a passive estimation setting to a setting with joint estimation and control. Section 5 presents a quantitative example. Section 6 relaxes the rational expectations assumption and considers the possibility of sticky and extrapolative expectations on the part of both real-world agents and agents inside the structural model. Section 6 also considers multivariate inference.

## 2. THE ECONOMIC SETTING

We initially consider a univariate parameter inference problem where the econometric model is exactly identified. The first subsection describes timing and technology assumptions. The second subsection illustrates how the general framework maps to a specific applied econometric problem.

*2.1. Timing, Technology, and Beliefs.* There is a real-world representative sample consisting of a continuum of atomistic agents (“firms”) privately endowed with a policy-invariant (“deep”) structural parameter. Knowledge of this parameter is sufficient for the government to set policy optimally.

An econometrician will observe an empirical moment derived from the measured actions of the sample firms. To fix ideas, one can think of the moment as being the sample mean of investment, new employees, R&D, or leverage. In practice, moments such as variance, skewness, or kurtosis may also be informative about deep firm-level parameters. In the context of indirect inference, the moment can be the coefficient obtained when firm decision variables are regressed on observable covariates, such as the coefficient on market-to-book ( $Q$ ) in an investment regression. The econometrician has developed a structural model and will match her model-implied moment with the observed empirical moment.

The real-world firms have rational expectations, and act noncooperatively. Each atomistic firm correctly understands it cannot change the moment observed by the econometrician by unilaterally changing its own action. The deep parameter, denoted as  $u$ , is common to all sample firms. However, this assumption does not preclude firm heterogeneity. For example, firms may be identical ex ante but face idiosyncratic shocks ex post. Alternatively, firms may face



idiosyncratic shocks that alter their measured actions. Finally, firm-level parameters might be, say, multiples of a common aggregate parameter  $u$ , for example,  $u_i = \varepsilon_i u$ , where  $\varepsilon_i$  is a firm-specific scalar known by firm  $i$ . An alternative technological assumption, not adopted here, is that each firm receives a noisy signal of the common parameter  $u$ . In such a setting, as in the present setting, parameter inference would need to account for feedback from inference to the policy variable.

The parameter  $u$  represents the realization of a random variable  $\tilde{u}$  with cumulative distribution function  $\Psi$  with a strictly positive density  $\psi$  on  $\mathbb{R}$  with no atoms. The realized parameter  $u$  is privately observed by each of the sample firms, but unobservable to the econometrician and the government. Below,  $\hat{u}(u)$  denotes an equilibrium parameter inference by the econometrician in the event that  $\tilde{u} = u$ , with  $\hat{u}(\cdot)$  denoting an equilibrium inference function.

Timing is as follows. When the model opens at time  $t = 0$ , the government policy variable is initially equal to the predetermined status quo  $\gamma_0 \in \Gamma$ , where the set of feasible government policies is  $\Gamma \equiv (\underline{\gamma}, \bar{\gamma})$ .<sup>8</sup> Next, nature draws  $u$  according to the distribution function  $\Psi$ . Each sample firm  $i$  then chooses an optimal *preference action*  $\phi_i$ . This action can be multidimensional. The econometrician then observes the *empirical moment*  $m$ , which is derived from the preference actions of the sample firms. Next, the econometrician will attempt to match her model-implied moment with the empirical moment, resulting in *parameter inference*  $\hat{u}$ . The econometrician then reports  $\hat{u}$  to the government. All of these events take place at the initial time  $t = 0$ .

Time is either discrete or continuous and the horizon can be finite or infinite. There is a known independent stochastic process  $d$  such that for all  $t \geq 0$ ,  $d_t \in \{0, 1\}$ . Let

$$(1) \quad t^* \equiv \inf_{t \geq 0} d_t = 1.$$

At time  $t^*$ , the government will enjoy a one-time opportunity to permanently re-set the policy variable, having already received the econometrician's report. At all prior dates, policy is fixed at the status quo  $\gamma_0$ . We assume this discretionary state is decision relevant for the real-world agents who form the expectation:

$$(2) \quad \mathbb{E}[t^*] < \infty.$$

In contrast, one can understand standard moment matching methods as implicitly assuming that real-world agents believe the government will never enjoy policy discretion.

The *equilibrium discretionary policy* is denoted  $\gamma^*$ . Under the stated assumptions, government policy postinference is a stochastic process  $\tilde{\gamma}_t$  with

$$(3) \quad \begin{aligned} t < t^* &\Rightarrow \tilde{\gamma}_t = \gamma_0, \\ t \geq t^* &\Rightarrow \tilde{\gamma}_t = \gamma^*. \end{aligned}$$

No sample firm receives any signal that is informative about  $\gamma^*$  aside from  $u$ . Thus, firm policy expectations are homogeneous. With this in mind, let  $\gamma$  denote the value of  $\gamma^*$  anticipated by the sample firms conditional upon their knowledge of  $u$ .

The optimal preference action of real-world firm  $i$  can be expressed as a function of the parameter  $u$ , anticipated discretionary policy  $\gamma$ , and the status quo policy:

$$(4) \quad \phi_i(u, \gamma; \gamma_0).$$

Above the subscript  $i$  captures idiosyncratic shocks.

<sup>8</sup> This could be the real line, for example.

It is assumed that observation of a continuum of sample firms is sufficient to ensure that any idiosyncratic shocks have no effect on the observed moment, so that  $m$  can be expressed as  $m(u, \gamma; \gamma_0)$ . For brevity, the constant  $\gamma_0$  capturing the status quo policy will be suppressed and the empirical moment will be represented by the following mapping:

$$(5) \quad m : \mathbb{R} \times \Gamma \rightarrow \mathbb{R}.$$

The first argument in the moment function  $m$  is the unknown parameter  $u \in \mathbb{R}$ . The second argument in the moment function is the discretionary government policy  $\gamma \in \Gamma$  as anticipated by the real-world agents.

The following assumption ensures the setting considered is seemingly-ideal.

**ASSUMPTION 1.** *The model-implied moment function is identical to the empirical moment function  $m : \mathbb{R} \times \Gamma \rightarrow \mathbb{R}$ . Moreover, for each  $\gamma \in \Gamma$ , the function  $m(\cdot, \gamma)$  is continuously differentiable and strictly monotonic.*

The first part of Assumption 1 states that, at least in terms of capturing the real technology, the structural model is correct. In particular, from Assumption 1 it follows that if the structural modeler were to correctly stipulate  $u$  and  $\gamma$ , the model-implied moment would match the real-world moment. Notice, the second part of Assumption 1 ensures that we satisfy the *traditional* moment matching identifying assumption that  $m(\cdot, \gamma)$  is strictly monotonic in the unknown parameter.

We next characterize how the moment varies with anticipated discretionary government policy.

**ASSUMPTION 2.** *For each  $u \in \mathbb{R}$ ,  $m(u, \cdot)$  is a continuously differentiable strictly monotonic function.*

Notice, the setting considered is quite general. For example, as in Blume et al. (1982), one can think of the sample firms as solving canonical finite or infinite horizon dynamic programming problems with differentiable policy functions where monotone comparative statics apply and carry over to  $m$ . Nevertheless, it is worth emphasizing that in order for Assumption 2 to hold, it must be the case that the sample firms are solving forward-looking problems in which anticipated discretionary government policy  $\gamma$  enters as a relevant parameter in their program, through periodic payoff functions, constraint functions, and/or transition functions.

The function  $g : \mathbb{R} \rightarrow \Gamma$  represents discretionary government policy. If the government had the ability to directly observe  $u$ , the true optimal discretionary policy would be  $g(u)$ . The following assumption is imposed:

**ASSUMPTION 3.** *Discretionary government policy  $g$  is a continuously differentiable strictly monotonic function mapping  $\mathbb{R}$  onto  $\Gamma$ .*

The government is presumed to believe that the standard moment matching exercise will allow the econometrician to deliver a correct estimate of the unknown parameter. Critically, Assumption 1 would seem to imply that this confidence is justified. After all, the model moment function is equal to the empirical moment function, and the moment is monotone in the unknown parameter. We have the following assumption:

**ASSUMPTION 4.** *The government chooses discretionary policy based upon a belief that for all  $u \in \mathbb{R}$ ,  $\widehat{u}(u) = u$ .*

From Assumption 4, it follows that for all  $u \in \mathbb{R}$ , the *endogenous discretionary policy* of the government is

$$(6) \quad \gamma^*(u) = g[\widehat{u}(u)].$$

An alternative interpretation of condition (6) is that the function  $g$  represents equilibrium policy outcomes from an extensive form game in which the econometrician's parameter estimate is fed into the political process. This alternative interpretation would not alter our characterization of bias in parameter inference, but would necessarily rule out discussions of policy bias.

Recall, we posit that the real-world agents have rational expectations. In particular, real-world agents know that the government may enjoy policy discretion at some future date. They also know the government will place full faith in the econometrician's structural parameter estimate  $\widehat{u}$  (Assumption 4), and will then input this estimate into the policy function  $g$  (Assumption 3). The following assumption formalizes this specification of beliefs:

**ASSUMPTION 5 (REAL-WORLD AGENT RATIONAL EXPECTATIONS).** *For all  $u \in \mathbb{R}$ , real-world firms correctly anticipate discretionary government policy, with*

$$(7) \quad \gamma(u) = \gamma^*(u) = g[\widehat{u}(u)].$$

The first equality in the preceding equation ensures that  $\gamma(\cdot)$  satisfies rational expectations. The second equality reflects how discretionary government policy  $\gamma^*$  will actually be formed in equilibrium, with  $\widehat{u}(u)$  being fed into  $g$ . Effectively, under rational expectations, the real-world firms infer the econometrician's parameter estimate, which allows them to correctly anticipate discretionary government policy.

From the preceding equation, it follows that the real-world empirical moment observed by the econometrician is:

$$(8) \quad m[u, \gamma(u)] = m[u, \gamma^*(u)] = m[u, g(\widehat{u}(u))].$$

Finally, we assume the econometrician departs from rational expectations by parameterizing her structural model according to the status quo. We have the following assumption:

**ASSUMPTION 6 (STATUS QUO MODEL PARAMETERIZATION).** *Firms inside the structural model incorrectly anticipate that the status quo will be maintained even if the government enjoys policy discretion, holding the incorrect belief that for all  $u$ ,  $\gamma(u) = \gamma_0$ .*

From the preceding discussion, it follows that for all  $u \in \mathbb{R}$ , the structural econometrician's parameter estimate will be derived from the following *inference equation*:

$$(9) \quad m[u, \gamma^*(u)] = m[\widehat{u}(u), \gamma_0]$$

or

$$(10) \quad m[u, g(\widehat{u}(u))] = m[\widehat{u}(u), \gamma_0].$$

The left side of the preceding equation is the real-world empirical moment. The empirical moment reflects the fact that the sample firms will choose their preinference actions optimally given the true parameter value  $u$  and their correct anticipation of discretionary government policy (Assumption 5). The right side of the preceding equation is the model-implied moment under the status quo parameterization (Assumption 6). The estimated parameter  $\widehat{u}(u)$  is chosen so that the model implied moment is equal to the observed empirical moment.

Before proceeding, it is worthwhile to consider an alternative, more complex, motivation for the inference equation (10) since this equation serves as the foundation for subsequent results regarding bias and bias correction. In particular, suppose instead the structural model does not treat government policy as fixed forever at  $\gamma_0$  but instead treats government policy as an exogenous stochastic process as in, say, Keane and Wolpin (2002). In order to approximate such an inference approach within our framework, one can think of the structural model as treating government policy as an independent discrete-state Markov chain with one state, call it state 0, being the state in which the government will enjoy policy discretion. The structural model then incorrectly treats government policy in state 0 as being an exogenous parameter  $\gamma_0$  whereas the real-world firms understand that government policy in the discretionary state 0 will be endogenously set at  $g(\hat{u}(u))$ . The inference equation (10) still applies in such a setting, and consequently, so do all the results that follow below. Having said this, it is clear that the approach of Keane and Wolpin (2002), while violating rational expectations in exercises of joint estimation and control, still offers an improvement over the common practice of treating policy as fixed forever at the status quo.

2.2. *Example: Inferring Labor Adjustment Costs.* At this stage it will be useful to fix ideas by considering a stripped-down example of the type of inference problem subsumed by our model. To this end, consider an econometrician who wants to estimate a labor adjustment cost parameter  $u$  based upon some empirical moment, say, the average change in firm or plant-level employment. This exercise is in the spirit of Hammermesh (1989), Blanchard and Portugal (2001), and Ejarque and Portugal (2007), who estimate parameters of labor adjustment cost functions and then use the estimates as the basis for making policy recommendations regarding labor market reforms. Although the focus of the example is on labor adjustment costs, similar arguments would apply to moment-based inference of capital stock adjustment cost parameters.

Let  $\phi_i$  denote the number of workers hired by firm  $i$ . Firms face quadratic costs of bringing new employees onto their workforce, with the costs increasing in units of governmental regulation. The status quo features  $\gamma_0$  units of regulation. The government will enjoy policy discretion with probability  $p > 0$  and real-world firms anticipate  $\gamma$  units of discretionary regulation. The sample firms make their hiring decisions before the policy uncertainty is resolved. Each real-world firm solves the following linear-quadratic program:

$$(11) \quad \max_{\phi_i} \phi_i q - \frac{1}{2} [p\gamma + (1 - p)\gamma_0] N(u) (\phi_i - \varepsilon_i)^2.$$

In the preceding equation,  $q > 0$  represents the shadow value of an “installed” worker—the net present value of marginal product less wages. For simplicity, assume  $q$  is known to the econometrician.<sup>9</sup> The function  $N$  is, say, the normal cumulative distribution. This function is scaled by expected units of regulation. The term  $\varepsilon_i$  is a mean-zero firm-specific shock. In this way, the structural estimation allows for heterogeneity.

Since the real-world firms have rational expectations, they anticipate discretionary government policy  $\gamma = \gamma^*$ , and the econometrician will observe the following empirical moment:

$$(12) \quad \int_i \phi_i di = m[u, \gamma^*(u)] = [pg(\hat{u}(u)) + (1 - p)\gamma_0]^{-1} [N(u)]^{-1} q.$$

The econometrician chooses her parameter estimate so that the model-implied moment is just equal to the observed empirical moment. The inference equation (10) is:

$$(13) \quad [pg(\hat{u}(u)) + (1 - p)\gamma_0]^{-1} [N(u)]^{-1} q = [\gamma_0]^{-1} [N(\hat{u}(u))]^{-1} q.$$

<sup>9</sup> Or the econometrician is willing to rely upon existing estimates of this parameter.

Rearranging terms in the preceding equation we find

$$(14) \quad \widehat{u}(u) = N^{-1} \left[ \frac{pg[\widehat{u}(u)] + (1-p)\gamma_0}{\gamma_0} \times N(u) \right].$$

From the preceding equation it follows that

$$(15) \quad \widehat{u}(u) = u \Leftrightarrow g(u) = \gamma_0.$$

Notice, the preceding equation implies that parameter inference is unbiased at point  $u$  if and only if the status quo is actually optimal at that point—a troubling implication. Anticipating, the next section shows this finding carries over more generally.

### 3. BIAS CHARACTERIZATION

This section characterizes the nature of parameter inference and associated policy outcomes if the structural model fails to impose the assumption that firms have rational expectations.

Before proceeding, it will be convenient to express the differential form of the inference equation. In particular, under technical conditions derived below, there will exist a continuously differentiable function  $\widehat{u}(\cdot)$  satisfying the inference equation (10). Assuming such a function exists, we have the following differential form:

$$(16) \quad m_u[u, g(\widehat{u}(u))] + m_\gamma[u, g(\widehat{u}(u))]g'[\widehat{u}(u)]\widehat{u}'(u) = m_u[\widehat{u}(u), \gamma_0]\widehat{u}'(u).$$

The differential form of the inference equation makes clear the potential for bias. The right side captures the econometrician's faulty inference procedure, which is predicated upon the incorrect assumption that firms expect the status quo to be maintained with probability 1. Thus, she incorrectly imputes any change in the observed moment to the direct effect as captured by the partial derivative,  $m_u$ . The left side of the preceding equation captures the true total differential of the empirical moment with respect to  $u$ . If  $u$  is perturbed, there will be a direct effect on the moment as captured by the first term,  $m_u$ . In addition, the empirical moment will vary due to the rational anticipation of firms that government policy will change based upon changes in the inferred parameter value. This inference-policy feedback effect is captured by the second term on the left side of the equation ( $m_\gamma g' \widehat{u}'$ ).

Let  $u_0$  be the unique value of the parameter  $u$  at which a fully informed government would find it optimal to implement the status quo policy  $\gamma_0$ . That is,

$$(17) \quad u_0 \equiv g^{-1}(\gamma_0) \Leftrightarrow g(u_0) = \gamma_0.$$

Uniqueness of  $u_0$  and invertibility follow from  $g$  being strictly monotone (Assumption 3).

The next proposition characterizes the realization(s) of the random variable  $\widehat{u}$  at which parameter inference will be unbiased.

**PROPOSITION 1.** *Let the structural model be parameterized assuming government will implement  $\gamma_0$  (the status quo) even if it enjoys policy discretion. Parameter inference is unbiased at point  $u$  if and only if  $g(u) = \gamma_0$ . There is a unique point at which this occurs,  $u_0 \equiv g^{-1}(\gamma_0)$ .*

**PROOF.** Referring to the inference equation (9), it follows from the strict monotonicity of  $m$  in its first argument that

$$\gamma^*(u) = \gamma_0 \Rightarrow \widehat{u}(u) = u.$$

Again referring to the inference equation (9), it follows from the strict monotonicity of  $m$  in its second argument that

$$\widehat{u}(u) = u \Rightarrow \gamma^*(u) = \gamma_0.$$

Finally if point  $u$  is a point such that parameter inference is unbiased and the status quo is optimal then it must be that

$$\gamma^*(u) = g(u) = \gamma_0.$$

From the strict monotonicity of  $g$  the unique point at which this occurs,  $u_0$ . □

The intuition for the preceding result is as follows. At any realization of  $u$  other than  $u_0$ , real-world firms anticipate the government will implement a policy different from the status quo should it enjoy policy discretion. The real-world firms then change their optimal behavior accordingly, leading to changes in the observed moment. However, under Assumption 6, the econometrician fails to take the inference-policy feedback effect into account, leading to bias.

Having established parameter inference will only be unbiased at point  $u_0$ , the next proposition provides insight into the nature of bias at all other  $u \in \mathbb{R}$ .

**PROPOSITION 2.** *Let the inference equation (9) be satisfied at point  $u$  by  $\widehat{u}(u)$ . If  $m_u m_\gamma > 0$ , then*

$$\begin{aligned} \gamma^*(u) < \gamma_0 &\Rightarrow \widehat{u}(u) < u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow \widehat{u}(u) > u. \end{aligned}$$

*If  $m_u m_\gamma < 0$ , then*

$$\begin{aligned} \gamma^*(u) < \gamma_0 &\Rightarrow \widehat{u}(u) > u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow \widehat{u}(u) < u. \end{aligned}$$

**PROOF.** There are four cases to consider. Suppose first  $m$  is increasing in both arguments. Then, from the inference equation (9) it follows

$$\begin{aligned} \gamma^*(u) < \gamma_0 &\Rightarrow m[u, \gamma_0] > m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) < u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow m[u, \gamma_0] < m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) > u. \end{aligned}$$

Suppose next  $m$  is decreasing in both arguments. Then,

$$\begin{aligned} \gamma^*(u) < \gamma_0 &\Rightarrow m[u, \gamma_0] < m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) < u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow m[u, \gamma_0] > m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) > u. \end{aligned}$$

Suppose next  $m$  is decreasing in its first argument and increasing in its second argument. Then,

$$\begin{aligned} \gamma^*(u) < \gamma_0 &\Rightarrow m[u, \gamma_0] > m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) > u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow m[u, \gamma_0] < m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) < u. \end{aligned}$$

Suppose finally  $m$  is increasing in its first argument and decreasing in its second argument. Then,

$$\begin{aligned}\gamma^*(u) < \gamma_0 &\Rightarrow m[u, \gamma_0] < m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) > u, \\ \gamma^*(u) > \gamma_0 &\Rightarrow m[u, \gamma_0] > m[u, \gamma^*(u)] \equiv m[\widehat{u}(u), \gamma_0] \Rightarrow \widehat{u}(u) < u.\end{aligned}$$

□

The intuition behind the preceding result is as follows. Taking the first part of the proposition, suppose the empirical moment function  $m$  is increasing (decreasing) in both arguments. Then if, say,  $\gamma^*(u) > \gamma_0$ , the moment will be higher (lower) than would be inferred based upon the direct effect  $m_u$ , causing  $\widehat{u}$  to overshoot  $u$ . Taking the second part of the proposition, suppose  $m_u > 0$  and  $m_\gamma < 0$ . Then if, say,  $\gamma^*(u) > \gamma_0$ , the moment will be lower than would be inferred based upon the direct effect  $m_u$ , causing  $\widehat{u}$  to undershoot  $u$ .

The preceding proposition characterizes  $\widehat{u}$  at a *particular point*  $u$  where the inference equation (9) has a solution. However, as shown below, the inference equation need not have a solution. With this in mind, the following lemma offers a sufficient condition for the existence of a (continuously differentiable) function  $\widehat{u}(\cdot)$  satisfying the inference equation pointwise for all  $u \in \mathbb{R}$ .

**LEMMA 1.** *Let  $m_u m_\gamma < 0$  and  $g' > 0$  or let  $m_u m_\gamma > 0$  and  $g' < 0$ . Then there exists a continuously differentiable strictly monotonic increasing function  $\widehat{u}(\cdot)$  satisfying the inference equation (9) for all  $u \in \mathbb{R}$ . The function  $\widehat{u}(\cdot)$  has slope in  $(0, 1)$  at  $u_0$ .*

**PROOF.** Consider the following function, which is continuously differentiable in its two arguments

$$F(u, z) \equiv m[u, g(z)] - m(z, \gamma_0).$$

Any root  $z$  of the preceding equation represents a solution to the inference equation (9). We know (Proposition 1) the root at  $u_0$  is  $u_0$ . Consider next arbitrary  $u \neq u_0$ . Under the stated conditions it is readily verified that

$$\begin{aligned}F(u, u) &\equiv m[u, g(u)] - m(u, \gamma_0) \\ F(u, u_0) &\equiv m(u, \gamma_0) - m(u_0, \gamma_0)\end{aligned}$$

have opposite signs. From the Location of Roots Theorem, there exists a point  $\widehat{u}$  solving the inference equation

$$F(u, \widehat{u}) = 0.$$

Moreover, under the stated conditions

$$\frac{\partial}{\partial \widehat{u}} F(u, \widehat{u}) = m_\gamma [u, g(\widehat{u})] g'(\widehat{u}) - m_u(\widehat{u}, \gamma_0) \neq 0.$$

It thus follows from the Implicit Function Theorem that there exists a continuously differentiable function  $\widehat{u}(\cdot)$  defined on an interval  $I$  about the (arbitrary) point  $u$  such that

$$(18) \quad F[\widetilde{u}, \widehat{u}(\widetilde{u})] = 0 \quad \forall \widetilde{u} \in I$$

and

$$\begin{aligned}
 (19) \quad \widehat{u}(u) &= \frac{m_u[u, g(\widehat{u}(u))]}{m_u[\widehat{u}(u), \gamma_0] - m_\gamma[u, g(\widehat{u}(u))]g'[\widehat{u}(u)]} \\
 &= \left[ \frac{m_u[\widehat{u}(u), \gamma_0]}{m_u[u, g(\widehat{u}(u))]} - \frac{m_\gamma[u, g(\widehat{u}(u))]g'[\widehat{u}(u)]}{m_u[u, g(\widehat{u}(u))]} \right]^{-1}.
 \end{aligned}$$

Notice, under the stated conditions, the term in square brackets in the preceding equation is strictly positive, implying the derivative of the function  $\widehat{u}$  is positive. Finally, the last statement in the lemma follows from

$$\begin{aligned}
 (20) \quad \widehat{u}'(u_0) &= \frac{m_u[u_0, g(\widehat{u}(u_0))]}{m_u[\widehat{u}(u_0), \gamma_0] - m_\gamma[u_0, g(\widehat{u}(u_0))]g'[\widehat{u}(u_0)]} \\
 &= \frac{m_u[u_0, g(u_0)]}{m_u(u_0, \gamma_0) - m_\gamma[u_0, g(u_0)]g'(u_0)} \\
 &= \left[ 1 - \frac{m_\gamma[u_0, g(u_0)]g'(u_0)}{m_u(u_0, \gamma_0)} \right]^{-1}.
 \end{aligned}$$

□

In order to illustrate the preceding lemma, and many that follow, it will be useful to define a *linear technology*:

$$\begin{aligned}
 (21) \quad m(u, \gamma) &\equiv \alpha u + \beta \gamma, \\
 g(\widehat{u}) &\equiv \kappa \widehat{u},
 \end{aligned}$$

where  $\alpha, \beta,$  and  $\kappa$  are arbitrary nonzero constants. Under the linear technology, the inference equation (10) can be written as

$$u + \kappa \widehat{u}(u) = \widehat{u}(u) + \gamma_0.$$

From Equation (17) it follows that here  $\gamma_0 = \kappa u_0$ . Using this fact, and rearranging terms in the preceding equation, the inference equation can be expressed as

$$(22) \quad \alpha u - \beta \kappa u_0 = (\alpha - \beta \kappa) \widehat{u}(u).$$

If  $\alpha = \beta \kappa$ , the preceding equation does not have a solution at any point other than  $u_0$ . Under the conditions in Lemma 1,  $\alpha \neq \beta \kappa$ . In fact, under the conditions specified in the lemma,  $\alpha$  and  $\beta \kappa$  have different signs. With  $\alpha \neq \beta \kappa$ , the solution to the linear technology inference equation is

$$(23) \quad \widehat{u}(u) = \frac{\alpha u - \beta \kappa u_0}{\alpha - \beta \kappa} = u + \frac{\beta \kappa (u - u_0)}{\alpha - \beta \kappa}.$$

Under the conditions stated in Lemma 1,  $\widehat{u}'$  is some constant in (0,1).

Lemma 1 leads directly to the following proposition.

**PROPOSITION 3.** *Let  $m_u m_\gamma < 0$  and  $g' > 0$  or let  $m_u m_\gamma > 0$  and  $g' < 0$ . Then there exists a continuously differentiable strictly monotonic increasing function  $\widehat{u}(\cdot)$  satisfying the inference equation. For all  $u < u_0, \widehat{u}(u) \in (u, u_0)$  and for all  $u > u_0, \widehat{u}(u) \in (u_0, u)$ . If  $g$  is increasing, then  $u < u_0$  implies  $\gamma^*(u) \in (g(u), \gamma_0)$  and  $u > u_0$  implies  $\gamma^*(u) \in (\gamma_0, g(u))$ . If  $g$  is decreasing, then  $u < u_0$  implies  $\gamma^*(u) \in (\gamma_0, g(u))$  and  $u > u_0$  implies  $\gamma^*(u) \in (g(u), \gamma_0)$ .*



PROOF. The first statement in the proposition is from Lemma 1. Next note that  $\widehat{u}(u_0) \in (0, 1)$ . It follows that for  $u$  on the left neighborhood of  $u_0$ ,  $\widehat{u}(u) \in (u, u_0)$  and for  $u$  on the right neighborhood of  $u_0$ ,  $\widehat{u}(u) \in (u_0, u)$ . From the continuity of  $\widehat{u}(\cdot)$  and Proposition 1 it follows that for all  $u < u_0$ ,  $\widehat{u}(u) > u$  and for all  $u > u_0$ ,  $\widehat{u}(u) < u$ . From the strict monotonicity of  $\widehat{u}(\cdot)$  it follows that for all  $u < u_0$ ,  $\widehat{u}(u) < u_0$  and for all  $u > u_0$ ,  $\widehat{u}(u) > u_0$ . The final two statements in the proposition follow from the fact that  $\gamma^* = g(\widehat{u})$ .  $\square$

Inspection of Equation (16) reveals the intuition for the preceding proposition. Under the stated assumptions, the second term on the left side of the differential form of the inference equation (16) dampens the sensitivity of the moment to changes in  $u$ —an effect ignored by the econometrician. She will then incorrectly impute the small changes in the moment to small changes in  $u$ . That is,  $\widehat{u}$  will tend to have a slope less than unity, with  $\widehat{u}$  overshooting for  $u < u_0$  and undershooting for  $u > u_0$ .

These effects are illustrated in Figures 1, 2, and 3, which consider the linear technology with  $m = u - \gamma$  and  $g = \widehat{u}/2$ , with  $u_0 = 0$ . Equation (23) pins down the inference function here, with  $\widehat{u}'(u) = 2/3$ . Figure 1 contrasts the true empirical moment function  $m[u, g(\widehat{u}(u))]$  and the econometrician's model-implied moment function  $m(u, \gamma_0)$ . The former accounts for policy feedback and the latter fails to do so. Here, the econometrician incorrectly imputes the dampened sensitivity of the observed moment to changes in  $u$  to small changes in  $u$ . Figure 2 shows the resulting single crossing of  $\widehat{u}$  with the 45 degree line from above, consistent with the notion of dampened sensitivity. Finally, since  $g$  has here been assumed to be increasing, Figure 3 shows the resulting policy overshooting relative to the optimal policy for low values of  $u$  and undershooting relative to the optimal policy for high values of  $u$ .

We next consider the nature of inference and policy bias under alternative technologies. However, before doing so, we must establish a sufficient condition for the existence of a well-behaved solution to the inference equation. After all, if we consider departures from the technologies assumed in the preceding proposition, it is possible that there is no solution to the inference equation. To see this, consider the linear technology and suppose that, departing from the preceding two propositions,  $\alpha$  and  $\beta$  have the same sign and  $\kappa > 0$  or  $\alpha$  and  $\beta$  have different signs and  $\kappa > 0$ . In either case, it is possible that  $\alpha = \beta\kappa$  so that there is no solution to the inference equation. With such a possibility in mind, the next lemma provides a sufficient condition for the existence of a continuously differentiable solution to the inference equation.

LEMMA 2. *If*

$$(24) \quad m_1(x, \gamma_0) \neq m_2[u, g(x)]g'(x) \quad \forall (x, u) \in \mathbb{R} \times \mathbb{R},$$

*then there exists a continuously differentiable strictly monotone function  $\widehat{u}(\cdot)$  satisfying the inference equation (9) for all  $u \in \mathbb{R}$ .*

PROOF. Define the following candidate solution to the inference equation:

$$\widehat{u}(u) \equiv u_0 + \int_{u_0}^u \frac{m_u[v, g(\widehat{u}(v))]}{m_u[\widehat{u}(v), \gamma_0] - m_\gamma[v, g(\widehat{u}(v))]g'[\widehat{u}(v)]} dv.$$

Since here  $\widehat{u}(u_0) = u_0$ , the candidate solution satisfies the inference equation at  $u_0$  (Proposition 1). Further, under the stated assumptions, the candidate solution has a well-defined derivative at all points, given in Equation (19). Rearranging terms in Equation (19), it follows that the candidate solution satisfies the differential form of the inference equation (16) pointwise. Thus,  $\widehat{u}$  is a continuous and differentiable solution to the inference equation. Moreover,  $\widehat{u}$  is continuously differentiable since  $m$  and  $g$  are continuously differentiable. Finally, the sign of the numerator in Equation (19) is constant. And, the sign of the denominator of this same equation cannot change since, by the Location of Roots Theorem, this would imply the

existence of an intermediate point such that the inequality in Equation (24) is violated. Thus,  $\widehat{u}$  must be strictly monotonic. □

To take a specific example, if the conditions of Lemma 2 were to be satisfied in the context of the linear technology (Equation (21)), then it follows  $\alpha \neq \beta\kappa$  and the linear technology inference function (23) along with its derivative would be well defined.

We have the following proposition:

**PROPOSITION 4.** *Let  $m_u m_\gamma > 0$  and  $g' > 0$  or let  $m_u m_\gamma < 0$  and  $g' < 0$ , with condition (24) being satisfied. If*

$$\frac{m_\gamma(u_0, \gamma_0)g'(u_0)}{m_u(u_0, \gamma_0)} < 1,$$

*there exists a continuously differentiable strictly monotonic increasing function  $\widehat{u}(\cdot)$  satisfying the inference equation. For all  $u < u_0$ ,  $\widehat{u}(u) < u$  and for all  $u > u_0$ ,  $\widehat{u}(u) > u$ . If  $g$  is increasing, then  $u < u_0$  implies  $\gamma^*(u) < g(u)$  and  $u > u_0$  implies  $\gamma^*(u) > g(u)$ . If  $g$  is decreasing, then  $u < u_0$  implies  $\gamma^*(u) > g(u)$  and  $u > u_0$  implies  $\gamma^*(u) < g(u)$ .*

*If*

$$\frac{m_\gamma(u_0, \gamma_0)g'(u_0)}{m_u(u_0, \gamma_0)} > 1,$$

*then there exists a continuously differentiable strictly monotonic decreasing function  $\widehat{u}(\cdot)$  satisfying the inference equation. For all  $u < u_0$ ,  $\widehat{u}(u) > u_0 > u$  and for all  $u > u_0$ ,  $\widehat{u}(u) < u_0 < u$ . If  $g$  is increasing then  $u < u_0$  implies  $\gamma^*(u) > \gamma_0 > g(u)$  and  $u > u_0$  implies  $\gamma^*(u) < \gamma_0 < g(u)$ . If  $g$  is decreasing, then  $u < u_0$  implies  $\gamma^*(u) < \gamma_0 < g(u)$  and  $u > u_0$  implies  $\gamma^*(u) > \gamma_0 > g(u)$ .*

**PROOF.** From Lemma 2 there exists a continuously differentiable strictly monotonic solution to the inference equation. From the final line in Equation (20) it follows

$$\frac{m_\gamma(u_0, \gamma_0)g'(u_0)}{m_u(u_0, \gamma_0)} < 1 \Rightarrow \widehat{u}'(u_0) > 1.$$

Considering this case,  $\widehat{u}$  must be strictly monotone increasing. Moreover, on the left neighborhood of  $u_0$ ,  $\widehat{u}(u) < u$  and on the right neighborhood of  $u_0$ ,  $\widehat{u}(u) > u$ . From the continuity of  $\widehat{u}$  and Proposition 1 it follows that for all  $u < u_0$ ,  $\widehat{u}(u) < u$  and for all  $u > u_0$ ,  $\widehat{u}(u) > u$ .

For the second part of the proposition, note that

$$\frac{m_\gamma(u_0, \gamma_0)g'(u_0)}{m_u(u_0, \gamma_0)} > 1 \Rightarrow \widehat{u}'(u_0) < 0.$$

Considering this case,  $\widehat{u}$  must be strictly monotone decreasing. It follows that for all  $u < u_0$ ,  $\widehat{u}(u) > u_0 > u$  and for all  $u > u_0$ ,  $\widehat{u}(u) < u_0 < u$ . The clauses pertaining to discretionary government policy follow from the fact that  $\gamma^* = g(\widehat{u})$ . □

Inspection of Equation (16) reveals the intuition for the first part of the preceding proposition. Under the posited technologies, the policy feedback effect causes the observed moment to be more sensitive to changes in  $u$  than is understood by the econometrician. She will then incorrectly impute large changes in the moment to large changes in  $u$ . That is,  $\widehat{u}$  will tend to have a slope in excess of unity, so that  $\widehat{u}$  undershoots for  $u < u_0$  and overshoots for  $u > u_0$ . In other words, the function  $\widehat{u}(u)$  will cross the function  $u$  at the point  $u_0$  from below.

These effects are illustrated in Figures 4, 5, and 6, which consider the linear technology with  $m = u + \gamma$  and  $g = \hat{u}/2$ , with  $u_0 = 0$ . Equation (23) pins down the inference function here, with  $\hat{u}(u) = 2$ . Figure 4 shows how the econometrician will incorrectly impute large changes in the moment to large changes in  $u$ . Figure 5 shows the resulting single crossing of  $\hat{u}$  with  $u$  from below. Finally, since  $g$  has here been assumed to be increasing, Figure 6 shows the resulting policy undershooting for low values of  $u$  and overshooting for high values of  $u$ .

The second part of the preceding proposition is illustrated most vividly by considering a particular example. To this end, consider the same linear moment  $m = u + \gamma$  but now assume  $g = 2\hat{u}$ , with  $u_0 = 0$ . That is, in the case being considered, discretionary government policy is more sensitive to the inferred value of the structural parameter. Equation (23) pins down the inference function here, with  $\hat{u}(u) = -u$  for all  $u$ . Notice, here we have a situation where the inferred value of the parameter has the wrong sign with probability 1. Of course, this implies that discretionary government policy will move in exactly the opposite direction relative to what is optimal.

Figures 7, 8, and 9 depict the nature of inference under this technology. For example, suppose the realized value is  $u = 5$ . Firms conjecture the econometrician will infer  $\hat{u}(5) = -5$  and anticipate discretionary governmental policy will be  $\gamma = 2\hat{u} = -10$ . The observed moment will be  $m = u + \gamma = 5 - 10 = -5$ . The econometrician incorrectly believes she is observing  $m = u + 2u_0 = u$  and so indeed draws the inference conjectured by the firms, with  $\hat{u} = -5$ . The government then implements  $\gamma^* = -10$ , consistent with the policy anticipated by the real-world firms.

#### 4. MOMENT MATCHING UNDER RATIONAL EXPECTATIONS

This section considers whether and how the econometrician can achieve unbiased parameter inference.

4.1. *Avoiding Bias and Achieving Optimality.* A natural to ask is whether it is possible to achieve unbiased parameter inference in the setting considered. Introspection suggests a ready solution. The underlying source of biased parameter inference in the preceding section was the failure of the econometrician to parameterize her model in a manner consistent with the rational expectations held by the firms (Assumption 6). Therefore, achieving unbiased inference would seem to necessitate “parameterizing” expectations correctly—with the issue being that the policy expectation is correctly understood as a function, instead of a parameter. Indeed, we have the following lemma:

LEMMA 3. *If firms anticipate monotone discretionary policy outcomes  $\gamma^{**}(\cdot)$ , then parameter inference will be unbiased for all  $u \in \mathbb{R}$  only if the structural model specifies discretionary policy outcomes as  $\gamma^{**}(\cdot)$ , with resulting rational expectations inference equation*

$$(25) \quad m[u, \gamma^{**}(u)] = m[\hat{u}(u), \gamma^{**}(u)].$$

PROOF. Suppose the structural model specifies firm beliefs according to some function  $\tilde{\gamma}(\cdot)$ . Then, the inference equation will be

$$(26) \quad m[u, \gamma^{**}(u)] = m[\hat{u}(u), \tilde{\gamma}(\hat{u}(u))].$$

Thus,

$$(27) \quad \hat{u}(u) = u \Rightarrow m[u, \gamma^{**}(u)] = m[u, \tilde{\gamma}(u)] \Rightarrow \tilde{\gamma} = \gamma^{**}.$$

The second implication follows from the strict monotonicity of  $m$  in its second argument.  $\square$

Of course, the government's ultimate objective is not to achieve unbiased parameter inference but rather to implement the optimal policy when it enjoys discretion. Therefore, the government would like to construct a rational expectations equilibrium predicated upon correct inference *and* firms anticipating a specific endogenous outcome

$$\gamma^{**}(\cdot) = g(\cdot).$$

But a necessary condition for correct parameter inference to be feasible for all  $u$  is that the empirical moment be invertible. To this end, let

$$(28) \quad \mu(u) \equiv m[u, g(u)].$$

We then have the following proposition:

**PROPOSITION 5.** *Let the empirical moment  $\mu(\cdot)$  (Equation (28)) be strictly monotone. Then, parameter inference will be unbiased for all  $u \in \mathbb{R}$  if and only if the structural model specifies discretionary policy outcomes as  $g(\cdot)$ .*

**PROOF.** The “only if” part of the proposition follows from Lemma 3. For sufficiency, suppose the structural model specifies firm beliefs according to some function  $\tilde{\gamma}(\cdot)$ . Then, the inference equation will be

$$(29) \quad m[u, g(u)] = m[\hat{u}(u), \tilde{\gamma}(\hat{u}(u))].$$

For sufficiency, note

$$\tilde{\gamma} = g \Rightarrow m[u, g(u)] = m[\hat{u}(u), g(\hat{u}(u))] \Rightarrow \hat{u}(u) = u.$$

□

It follows that in order for the econometrician to avoid bias and achieve first-best, she must replace the faulty inference equation (9) with the rational expectations inference equation

$$(30) \quad m[u, g(u)] = m[\hat{u}(u), g(u)].$$

Of course, the measured agents must understand the econometrician's procedure. Formally, in a rational expectations equilibrium there is no need for any agent to make a speech. Nevertheless, heuristically, in support of the postulated equilibrium, the econometrician could be understood as making the following speech to the firms.

I the structural econometrician will correctly infer the true value of the parameter  $u$  from the observation of the moment  $m$  that your actions generate. Further, armed with my correct inference, the government will implement the optimal policy  $g(u)$  should it enjoy policy discretion. And now that I have made this speech to you, I know that you know I will do this, and so you should anticipate  $g(u)$  as the discretionary government policy and, thus, act accordingly.

**4.2. Gallant and Tauchen Revisited.** In the title to their important paper, Gallant and Tauchen (1996) pose a question often asked by structural modelers: “Which Moment to Match?” An overarching message of our article is that moment-based inference changes if one is attempting joint estimation and control.

To illustrate, consider an econometrician operating in a world with linear technologies, with two competing moments being considered candidates for matching. In particular, suppose the

optimal government policy is  $\kappa u$ , where moments 1 and 2 have the following forms, respectively:

$$\begin{aligned} m_1 &\equiv \beta_1 \gamma, \\ m_2 &\equiv \alpha_2 u + \beta_2 \gamma, \\ \alpha_2 &\equiv -\beta_2 \kappa. \end{aligned}$$

According to the traditional moment selection criteria, moment 1 would be discarded since it violates the standard moment monotonicity condition (Assumption 1). In particular, according to the traditional moment selection criteria, moment 1 would be viewed as completely uninformative about the unknown parameter. In contrast, moment 2 would be viewed as informative about the unknown parameter.

But recall, the econometrician is engaged in an exercise of joint estimation and control, with the government attempting to achieve first-best. In this context, moment 1 is informative and moment 2 is uninformative. In order to see this, consider a conjectured rational expectations equilibrium with correct inference and first-best policy implementation. In such an equilibrium the two moments can be expressed as univariate functions of the unknown parameter. We have

$$(31) \quad \begin{aligned} \mu_1 &= \beta_1 \gamma^{**}(u) = \beta_1 g(u) = \beta_1 \kappa u, \\ \mu_2 &= \alpha_2 u + \beta_2 \gamma^{**}(u) = \alpha_2 u + \beta_2 g(u) = [\alpha_2 + \beta_2 \kappa] u = 0. \end{aligned}$$

Notice, we have here a situation where without policy feedback, moment 2 is informative and moment 1 is uninformative. Conversely, with policy feedback, moment 2 is uninformative and moment 1 is informative. Strikingly, moment 2 can be highly informative about the true value of the unknown parameter solely due to its sensitivity to the governmental policy variable. Intuitively, as  $u$  changes, so too does governmental policy in equilibrium, and this causes firm behavior to change in a manner informative about  $u$ .

We thus have the following proposition:

**PROPOSITION 6.** *Monotonicity of the moment function  $m(\cdot, \gamma)$  is neither necessary nor sufficient for  $m$  to be informative about the unknown parameter with joint estimation of  $u$  and control of  $\gamma^*$ .*

**4.3. An Algorithmic Approach to Moment-Based Inference.** The objective of this section is to propose a practically feasible algorithm allowing the econometrician to iterate to (approximately) correct inference of  $u$ , leading to a rational expectations equilibrium in which policy approximates first-best, with  $\gamma^*(u)$  arbitrarily close to  $g(u)$ .

To this end, consider the following *Algorithmic Inference Approach*:

- Start iteration  $n \in \{1, 2, 3, \dots\}$  with a provisional policy recommendation  $\gamma_n$ ;
- Let the iteration  $n$  inference  $\hat{u}_n$  solve

$$(32) \quad m_{observed} = m(\hat{u}_n, \gamma_n);$$

- Recompute the provisional government policy as  $\gamma_{n+1} = g(\hat{u}_n)$ ;
- Iterate until (approximate) internal consistency,  $|\gamma_{n+1} - \gamma_n| < \epsilon$  for  $\epsilon$  arbitrarily small.

We then have the following proposition showing that if  $\mu$  (Equation (28)) is strictly monotonic, elimination of internal inconsistency is sufficient to ensure correct inference and optimal government policy.

PROPOSITION 7. *Let  $\mu$  (Equation (28)) be strictly monotonic. At the  $n$ -th iteration, let the structural model be parameterized assuming government will implement  $\gamma_n$  should it enjoy policy discretion. The resulting inference  $\hat{u}_n$  will be equal to the true parameter  $u$  if and only if  $\hat{u}_n$  rationalizes  $\gamma_n$  so that policy convergence obtains with  $\gamma_n = g(\hat{u}_n) \equiv \gamma_{n+1}$ .*

PROOF. In order to establish sufficiency suppose  $\gamma_n = g(\hat{u}_n)$ . Under the stated conditions, the inference equation (9) can be rewritten as

$$m_{observed} = m[\hat{u}_n, g(\hat{u}_n)] \Rightarrow m_{observed} = \mu(\hat{u}_n).$$

From monotonicity of  $\mu$ , the unique value at which the observed moment matches the model-implied moment is the true  $u$ . In order to establish necessity, suppose  $\gamma_n \neq g(\hat{u}_n)$ . It then follows from the moment matching equation and monotonicity of  $m$  in its second argument that

$$m_{observed} = m(\hat{u}_n, \gamma_n) \neq m[\hat{u}_n, g(\hat{u}_n)] \equiv \mu(\hat{u}_n).$$

Since  $m_{observed} \neq \mu(\hat{u}_n)$  it follows  $\hat{u}_n \neq u$ . □

Of course, in practice, iteration will generally continue until approximate convergence. Therefore, it is interesting to evaluate the convergence properties of the preceding algorithm. Instead of doing so numerically with arbitrary examples, we first consider below iterating on the preceding algorithm in the case of the linear technology. To begin, note that iterating on  $\gamma_n$  values is equivalent to iterating on the  $u$  values that would justify them, for example,  $\kappa u_{n+1} \equiv \gamma_{n+1}$ .

In the posited rational expectations equilibrium, with first-best policy conjectured by the firms, the inference equation at iteration  $n + 1$  is

$$(33) \quad m[u, g(u)] = m[\hat{u}_{n+1}, \gamma_{n+1}].$$

With the linear technology, the preceding equation can be expressed as follows:

$$(34) \quad \alpha u + \beta \kappa u = \alpha \hat{u}_{n+1} + \beta \kappa \hat{u}_n.$$

Iterating on the preceding equation we have the following lemma, which shows that the proposed algorithm will converge to the truth provided the policy feedback effect is sufficiently weak relative to the direct effect.

LEMMA 4. *Under the linear technology (Equation (21)), the Algorithmic Inference Approach yields inference at the  $n$ -th iteration equal to*

$$(35) \quad \hat{u}_n = u + \left(-\frac{\beta \kappa}{\alpha}\right)^n (u_1 - u).$$

*The algorithm converges to the true parameter  $u$  for all  $u \in \mathbb{R}$  for all starting points  $u_1 \in \mathbb{R}$  if and only if*

$$\left| \frac{\beta \kappa}{\alpha} \right| < 1.$$

In fact, Lemma 4 is a special case of a more general convergence condition, which relies on bounding the policy feedback effect, as we show next.

PROPOSITION 8. *The Algorithmic Inference converges to the true parameter  $u$  for all  $u \in \mathbb{R}$  for all starting points  $\gamma_1 \in \Gamma$  if*

$$\left| \frac{m_\gamma g'}{m_u} \right| < 1.$$

PROOF. The inference equation is

$$m[u, g(u)] - m[\hat{u}_n, \gamma_n] = 0.$$

The preceding equation can be rewritten as

$$\{m[u, g(u)] - m[\hat{u}_n, g(u)]\} + \{m[\hat{u}_n, g(u)] - m[\hat{u}_n, \gamma_n]\} = 0.$$

From the mean value theorem, for each iteration  $n$ , there exists  $x_n$  between  $\hat{u}_n$  and  $u$ , and there exists  $g_n$  between  $g(u)$  and  $\gamma_n$  such that

$$m_u[x_n, g(u)](u - \hat{u}_n) + m_\gamma(\hat{u}_n, g_n)[g(u) - g(\hat{u}_{n-1})] = 0.$$

Applying the mean value theorem to the final term in the preceding equation, we know that for each iteration  $n$  there exists  $z_n$  between  $u$  and  $\hat{u}_{n-1}$  such that

$$m_u[x_n, g(u)](u - \hat{u}_n) + m_\gamma(\hat{u}_n, g_n)g'(z_n)(u - \hat{u}_{n-1}) = 0.$$

Rearranging terms in the preceding equation, we find that at each iteration  $n$

$$u - \hat{u}_n = -\frac{m_\gamma[\hat{u}_n, g_n]g'(z_n)}{m_u[x_n, g(u)]}(u - \hat{u}_{n-1}).$$

Under the stated condition  $\hat{u}_n$  converges to  $u$ . □

## 5. BIAS: A QUANTITATIVE EXAMPLE

This section considers an econometrician seeking to estimate unobserved costs of bankruptcy based upon the financial policies adopted by corporations. Understanding the magnitude of bankruptcy costs is important for a number of reasons. First, to the extent that bankruptcy costs are deadweight losses, instead of transfers, their magnitude is directly relevant for assessing the efficiency costs of corporate leverage, as well as tax-induced leverage increases. For example, in making the case for the Bush Administration Treasury for integration of the individual and corporate tax systems, Hubbard (1993) contended, “tax-induced distortions in corporations’ comparisons of nontax advantages and disadvantages of debt entail significant efficiency costs.” Second, the magnitude of bankruptcy costs is indirectly relevant to the tax authority estimating revenues. After all, higher bankruptcy costs serve as a counterweight to tax benefits of debt, discouraging firms from taking on extremely high leverage. For example, Gruber and Rauh (2007) estimate the tax elasticity of corporate income is only  $-0.2$ , evidence that would appear to contradict Hubbard’s notion that corporations aggressively change capital structures in response to tax incentives.

Early models, such as that of Stiglitz (1973), failed to deliver interior optimal leverage ratios. Lacking interior optimal leverage ratios, computational general equilibrium (CGE) models, for example, Ballard, et al. (1985), posited exogenous financing rules. In the absence of closed models, public finance economists such as Gordon and MacKie-Mason (1990) and Nadeau (1993) were forced into positing ad hoc costs of financial distress. In an important contribution, Leland (1994) showed how to develop a tractable logically closed model of

capital structure for firms facing taxation and costs of distress using contingent-claims pricing methods.

With the Leland model in mind, consider a structural econometrician who will observe the financing policies adopted by a set of homogeneous firms funding new investments during the preinference stage.<sup>10</sup> Specifically, the econometrician will measure the mean interest coverage ratio, as measured by the ratio of earnings before interest and taxes (EBIT) to interest expense. As shown below, this moment is directly informative about bankruptcy costs.

Consider first the decision problem of the firms. Each firm will choose a promised instantaneous coupon on a consol bond, denoted  $\phi$ . The firm will use the debt proceeds plus equity injections to fund a new investment, as is standard in project finance settings. We assume parameters are such that the investment has positive net present value. Formally, the new investment has positive net present value if the value of the levered enterprise exceeds the cost of the investment.

Debt enjoys a tax advantage, with interest being a deductible expense on the corporate income tax return. Consequently, each instant it is alive, the project firm will capture a gross tax shield equal to  $\phi\tilde{\gamma}$ , with the variable  $\tilde{\gamma}$  representing the corporate income tax rate that will be implemented just after the econometrician completes her parameter inference. The firm must weigh this debt tax shield benefit against costs of financial distress. In particular, in the event of EBIT being insufficient to service the coupon, the firm's debt will be canceled and bondholders will recover the unlevered firm value net of deadweight bankruptcy costs representing a fraction  $N(u)$  of unlevered firm value. The function  $N$  here is the standard normal cumulative distribution function.

Suppose firm EBIT follows a geometric Brownian motion with drift  $\mu$ , volatility  $\sigma$ , and initial value normalized at 1. The risk-free rate is denoted as  $r$ . The objective is to maximize levered project value. Or equivalently, firms maximize expected tax shield value minus expected default costs. Letting  $\gamma$  represent the anticipated tax rate, firms solve the following program:

$$(36) \quad \max_{\phi} \gamma \phi \left( \frac{1}{r} \right) (1 - \phi^{-\lambda}) - N(u) \frac{\phi(1 - \gamma)}{r - \mu} \phi^{-\lambda},$$

where  $\lambda$  is the negative root of the following quadratic equation:

$$(37) \quad \frac{1}{2} \sigma^2 \lambda^2 + \left( \mu - \frac{1}{2} \sigma^2 \right) \lambda - r = 0.$$

Note, the first term in the objective function captures tax shield value and the second term captures bankruptcy costs. Effectively, the tax shield represents an annuity that expires at the first passage of EBIT to the coupon from above. At this same point in time, bankruptcy costs incurred. This explains the presence of the term  $\phi^{-\lambda}$  in the objective function, which measures the price at date zero of a primitive claim paying 1 the first passage of EBIT to the coupon from above.

The first-order condition for the optimal coupon entails equating marginal tax benefits with marginal bankruptcy costs. In particular, the optimal coupon satisfies

$$(38) \quad \left( \frac{\gamma}{r} \right) [1 - (1 - \lambda) \phi^{-\lambda}] = (1 - \lambda) N(u) \frac{(1 - \gamma)}{r - \mu} \phi^{-\lambda}.$$

Rearranging terms in the preceding equation, it follows the optimal coupon is

$$(39) \quad \phi^* = (1 - \lambda)^{1/\lambda} \left[ 1 + N(u) \frac{(1 - \gamma)}{r - \mu} \frac{r}{\gamma} \right]^{1/\lambda}.$$

<sup>10</sup> The optimal coupon is linear in earnings before interest and taxes (EBIT), so coverage ratios will be equal if EBIT levels differ.



The moment observed by the econometrician, the mean interest coverage ratio, is  $1/\phi^*$ . Thus, in the present setting

$$(40) \quad m(u, \gamma) \equiv \mathbb{E}[\phi^{-1}] = (1 - \lambda)^{-1/\lambda} \left[ 1 + N(u) \frac{(1 - \gamma) r}{r - \mu \gamma} \right]^{-1/\lambda}.$$

Notice, in this particular case,  $m_u(u, \gamma) > 0$  and  $m_\gamma(u, \gamma) < 0$ . That is, the optimal interest coverage ratio is increasing in bankruptcy costs and decreasing in the tax rate.

Suppose now that the structural econometrician, who recommended the Trump tax cut, failed to impose the assumption that firms have rational expectations ( $\gamma = \gamma^*$ ). Specifically, suppose the econometrician treated the tax change as a counterfactual event and parameterized the model using the status quo tax rate. In the present context, the inference equation (9) takes the form

$$(41) \quad \begin{aligned} m[u, \gamma^*(u)] &= (1 - \lambda)^{-1/\lambda} \left[ 1 + N(u) \frac{[1 - \gamma^*(u)] r}{r - \mu \gamma^*(u)} \right]^{-1/\lambda} \\ &= (1 - \lambda)^{-1/\lambda} \left[ 1 + N(\hat{u}) \frac{(1 - \gamma_0) r}{r - \mu \gamma_0} \right]^{-1/\lambda} = m(\hat{u}, \gamma_0). \end{aligned}$$

Canceling terms in the preceding equation and rearranging terms one obtains:

$$(42) \quad N(\hat{u}) = \frac{[1 - \gamma^*(u)]/\gamma^*(u)}{(1 - \gamma_0)/\gamma_0} \times N(u).$$

How important quantitatively is the bias implied by the preceding equation? Following Goldstein et al. (2001), we can approximate the effect of personal taxes by setting  $\gamma_0$  and  $\gamma^*$  based upon the Miller (1977) debt tax shield formula. In particular, let  $\gamma_c$  denote the corporate tax rate,  $\gamma_e$  denote the equityholder tax rate, and  $\gamma_d$  denote the debtholder tax rate. The Miller debt tax shield value is

$$(43) \quad \gamma = 1 - \frac{(1 - \gamma_c)(1 - \gamma_e)}{(1 - \gamma_d)}.$$

Goldstein et al. (2001) estimate  $\gamma_c = 35\%$ ,  $\gamma_e = 20\%$ , and  $\gamma_d = 35\%$ . These parameter values are reflective of the status quo before the Trump corporate tax cut, which implies the status quo policy value is  $\gamma_0 = 20\%$ . The Trump tax reform cut the corporate income tax rate to  $\gamma_c = 21\%$ . This tax rate reduction substantially lowered the effective debt tax shield to  $\gamma^* = 2.8\%$ . Substituting these values into the bias formula in Equation (42) we find

$$N(\hat{u}) = 8.68 \times N(u).$$

That is, estimated bankruptcy costs here are 8.68 times actual bankruptcy costs. Intuitively, here the firms choose low leverage in rational anticipation of the upcoming tax cut. The econometrician treats the firms as ignorant of the prospective tax cut and treats the low leverage as indicative of very high bankruptcy costs.

## 6. EXTENSIONS

*6.1. Beyond Rational Expectations.* The preceding sections assumed real-world firms had fully rational expectations whereas firms inside the model were completely myopic. In this subsection, we relax these assumptions.

To begin, we assume instead that real-world agents place weight  $\omega > 0$  on  $\gamma^*(u)$  and weight  $(1 - \omega)$  on  $\gamma_0$ . A weight  $\omega < 1$  can be viewed as allowing for sticky expectations, whereas a weight  $\omega > 1$  may reflect extrapolative expectations.<sup>11</sup> Assuming linear technologies for ease of exposition, we have the following assumption:

ASSUMPTION 5' (REAL-WORLD AGENT EXPECTATIONS). *For all  $u \in \mathbb{R}$ , real-world firms anticipate government policy will be*

$$(44) \quad \gamma_a(u) = \kappa[\omega\widehat{u}(u) + (1 - \omega)u_0].$$

Next, we relax the assumption that firms inside the structural model are completely myopic. We have the following assumption:

ASSUMPTION 6' (MODEL PARAMETERIZATION). *Firms inside the structural model place weight  $w > 0$  on the true discretionary policy  $\gamma(u)$  and weight  $(1 - w)$  on  $\gamma_0$ , implying they anticipate discretionary government policy*

$$(45) \quad \gamma_e(u) = \kappa[w\widehat{u}(u) + (1 - w)u_0].$$

Under the stated assumptions the inference equation is

$$(46) \quad \alpha u + \beta\kappa\omega\widehat{u}(u) + \beta\kappa(1 - \omega)u_0 = \alpha\widehat{u}(u) + \beta\kappa w\widehat{u}(u) + \beta\kappa(1 - w)u_0.$$

The left side of the preceding equation is the real-world moment and the right side is the model-implied moment. From the preceding equation it follows that the inferred value of the parameter will be

$$(47) \quad \widehat{u}(u) \equiv u + \frac{[\beta\kappa(\omega - w)]}{[\beta\kappa(\omega - w) - \alpha]}(u_0 - u).$$

Notice, the preceding equation implies that even if we depart from rational expectations, there will be no bias so long as agents inside the model are hard-wired in the same way as the real-world agents, with  $w = \omega$ .

In the interest of brevity, we confine attention to the cases where  $\omega > w \geq 0$ . Then, overshooting and undershooting will obtain in the same parameter regions as in our main model.<sup>12</sup> If  $\alpha/\beta\kappa < 0$  and  $0 < \omega - w < 1$ ,  $|\widehat{u}(u) - u|$  is lower than in our main model for all parameter values and will increase with  $(\omega - w)$ . In fact, the preceding analysis subsumes our main model as a special case in which  $\omega = 1$  and  $w = 0$ .

We can also write a modified version of our iterative algorithm whereby at each step  $n$

$$(48) \quad \begin{aligned} \widehat{u}_n(u) &\equiv u + \frac{[\beta\kappa(\omega - w)]}{[\beta\kappa(\omega - w) - \alpha]}(u_0 - \widehat{u}_{n-1}) \\ &= u + \left[1 - \frac{\alpha}{[\beta\kappa(\omega - w)]}\right]^{-1} (u_0 - \widehat{u}_{n-1}). \end{aligned}$$

Then the following modified lemma obtains.

<sup>11</sup> On sticky and extrapolative expectations, see Enthoven and Arrow (1956), and Bouchaud et al. (2019).

<sup>12</sup> If  $\omega < w$ , the econometrician will assume agents place higher expectations on future discretionary policy than they actually do. Overshooting (undershooting) will obtain in parameter regions where there is undershooting (overshooting) in our main model.

LEMMA 4'. *Under the linear technology, the Algorithmic Inference Approach yields inference at the  $n$ -th iteration equal to*

$$(49) \quad \widehat{u}_n = u + \left( -\frac{\beta\kappa(\omega - w)}{\alpha} \right)^n (u_1 - u).$$

*The algorithm converges to the true parameter  $u$  for all  $u \in \mathbb{R}$  for all starting points  $u_1 \in \mathbb{R}$  if and only if*

$$\left| \frac{\beta\kappa(\omega - w)}{\alpha} \right| < 1.$$

6.2. *Multivariate Extension.* The preceding sections considered an econometrician attempting to infer one unknown parameter, with the government controlling one policy variable. In this section, we consider a multivariate extension. For simplicity, linearity is assumed.

There are  $n_u \geq 1$  unknown deep parameters, each with support on the real line. The realized vector is denoted  $\mathbf{u}$ . The econometrician seeks to infer  $\mathbf{u}$  based upon a vector  $\mathbf{m}$  consisting of  $n_u$  empirical moments. The government has  $n_\gamma \geq 1$  policy tools, with the full-information optimal policy being  $\mathbf{g}(\mathbf{u})$ .

The observed empirical moments are linear:

$$\mathbf{m} \equiv \mathbf{A}\mathbf{u} + \mathbf{B}\boldsymbol{\gamma}.$$

In the preceding equation,  $\mathbf{A}$  is an  $n_u \times n_u$  matrix of full rank with element  $\alpha_{ij}$  denoting the moment  $i$  coefficient on parameter  $u_j$ . Matrix  $\mathbf{B}$  is an  $n_u \times n_\gamma$  matrix with element  $\beta_{ij}$  denoting the moment  $i$  coefficient on government policy variable  $\gamma_j$ . The government policy vector is:

$$\boldsymbol{\gamma} = \mathbf{K}\widehat{\mathbf{u}}.$$

In the preceding equation,  $\mathbf{K}$  is an  $n_\gamma \times n_u$  matrix, with element  $\kappa_{ij}$  denoting the policy  $i$  coefficient on  $\widehat{u}_j$ .

Consider again the nature of bias that arises if the econometrician parameterizes government policy at the status quo

$$\boldsymbol{\gamma}_0 \equiv \mathbf{K}\mathbf{u}_0.$$

The inference equation is

$$(50) \quad \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{K}\widehat{\mathbf{u}} = \mathbf{A}\widehat{\mathbf{u}} + \mathbf{B}\mathbf{K}\mathbf{u}_0.$$

The left side of the preceding equation is the observed moment assuming real-world firms have rational expectations and the right side is the model-implied moment. Solving the preceding equation we obtain the multivariate analog of Equation (23):

$$(51) \quad \begin{aligned} \widehat{\mathbf{u}} &= \mathbf{u} + [\mathbf{A} - \mathbf{B}\mathbf{K}]^{-1} \mathbf{B}\mathbf{K}[\mathbf{u} - \mathbf{u}_0] \\ &= \mathbf{u} + [\mathbf{A} - \mathbf{B}\mathbf{K}]^{-1} \mathbf{B}[\mathbf{g}(\mathbf{u}) - \boldsymbol{\gamma}_0]. \end{aligned}$$

From the preceding equation it follows

$$(52) \quad \mathbf{u} = \mathbf{u}_0 \Rightarrow \mathbf{g}(\mathbf{u}) = \boldsymbol{\gamma}_0 \Rightarrow \widehat{\mathbf{u}} = \mathbf{u}.$$

It follows from the preceding equation that in the multivariate setting  $\mathbf{u} = \mathbf{u}_0$  is sufficient for absence of bias, but is not necessary. This is in contrast to the univariate case (Proposition 1) where  $u = u_0$  was both necessary and sufficient for absence of bias.

Other implications of the linear multivariate bias equation (51) are most readily illustrated by considering the simplest case with two unknown parameters and one government policy variable. In this case, let  $\beta_i$  denote the moment  $i$  coefficient on the government policy variable and let  $\kappa_j$  denote the government policy coefficient on  $\hat{u}_j$ . Applying Equation (51) we obtain:

$$(53) \quad \begin{aligned} \hat{u}_1 &= u_1 + \frac{[\beta_1 - \beta_2\alpha_{12}/\alpha_{22}][\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{11} - \beta_1\kappa_1 + [\beta_2\kappa_1\alpha_{12} + \beta_1\kappa_2\alpha_{21} - \beta_2\kappa_2\alpha_{11} - \alpha_{12}\alpha_{21}]/\alpha_{22}}, \\ \hat{u}_2 &= u_2 + \frac{[\beta_2 - \beta_1\alpha_{21}/\alpha_{11}][\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{22} - \beta_2\kappa_2 + [\beta_2\kappa_1\alpha_{12} + \beta_1\kappa_2\alpha_{21} - \beta_1\kappa_1\alpha_{22} - \alpha_{12}\alpha_{21}]/\alpha_{11}}. \end{aligned}$$

With the preceding equation in mind, suppose  $\alpha_{12} = \alpha_{21} = 0$ . That is, the moment  $i$  coefficient on parameter  $u_j$  is 0 for  $i \neq j$ . Here the traditional Jacobian formulation would suggest that the problem of inferring  $u_1$  is separable from the problem of inferring  $u_2$ . However, with policy feedback, it is apparent that the inference problems and biases are not separable, since

$$(54) \quad \begin{aligned} \hat{u}_1 &= u_1 + \frac{\beta_1[\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{11} - \beta_1\kappa_1 - \beta_2\kappa_2/\alpha_{22}}, \\ \hat{u}_2 &= u_2 + \frac{\beta_2[\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{22} - \beta_2\kappa_2 - \beta_1\kappa_1\alpha_{22}/\alpha_{11}}. \end{aligned}$$

Recall also that in the case of one unknown parameter, bias vanishes if: discretionary government policy is not affected by the econometrician’s estimate of the parameter ( $\kappa = 0$ ) or the moment is not affected by government policy ( $\beta = 0$ ), as shown in Equation (23). However, neither of these two conditions is sufficient to eliminate bias in a multivariate setting. In order to see this, consider again  $\alpha_{12} = \alpha_{21} = 0$ , and suppose also that the government policy variable does not depend upon  $\hat{u}_2$ , with  $\kappa_2 = 0$ . We then have

$$(55) \quad \begin{aligned} \hat{u}_1 &= u_1 + \frac{\beta_1\kappa_1(u_1 - u_{10})}{\alpha_{11} - \beta_1\kappa_1}, \\ \hat{u}_2 &= u_2 + \frac{\beta_2\kappa_1(u_1 - u_{10})}{\alpha_{22} - \beta_1\kappa_1\alpha_{22}/\alpha_{11}}. \end{aligned}$$

From the preceding equation it is apparent that even though  $\hat{u}_2$  does not inform policy,  $\hat{u}_2$  will nevertheless be biased so long as the government policy variable influences ( $\beta_2 \neq 0$ ) the respective moment (here  $m_2$ ) that is relied upon for inferring  $u_2$ .

It is also apparent that, in general, the existence of a moment that is independent of the government policy variable does not imply the absence of bias in any particular parameter estimate. In order to see this, suppose all four elements of matrix  $\mathbf{A}$  are positive. Suppose further that the government policy variable has no effect on one of the moments, say  $m_2$ , with  $\beta_2 = 0$ . In this case, bias still emerges, with

$$(56) \quad \begin{aligned} \hat{u}_1 &= u_1 + \frac{\beta_1[\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{11} - \beta_1\kappa_1 + [\beta_1\kappa_2\alpha_{21} - \alpha_{12}\alpha_{21}]/\alpha_{22}}, \\ \hat{u}_2 &= u_2 + \frac{-[\beta_1\alpha_{21}/\alpha_{11}][\kappa_1(u_1 - u_{10}) + \kappa_2(u_2 - u_{20})]}{\alpha_{22} + [\beta_1\kappa_2\alpha_{21} - \beta_1\kappa_1\alpha_{22} - \alpha_{12}\alpha_{21}]/\alpha_{11}}. \end{aligned}$$

Despite the subtle differences in the nature of bias arising in the univariate and multivariate cases, the solution of the problem is the same: consistent application of rational expectations. In order to see this, suppose now that the econometrician parameterizes the structural model in a manner consistent with the policies being recommended, with recommended policy  $\mathbf{K}\hat{\mathbf{u}}$  replacing the status quo policy  $\mathbf{K}\mathbf{u}_0$  in the original faulty inference equation (50). The rational expectations inference equation is:

$$(57) \quad \mathbf{A}\mathbf{u} + \mathbf{B}\mathbf{K}\hat{\mathbf{u}} = \mathbf{A}\hat{\mathbf{u}} + \mathbf{B}\mathbf{K}\hat{\mathbf{u}} \Rightarrow \hat{\mathbf{u}} = \mathbf{u}.$$

In order to see how this outcome can be achieved, consider the following extension of the algorithm we presented in the preceding section. Denote an  $n_u \times n_u$  matrix

$$\mathbf{M} = [\mathbf{A} - \mathbf{B}\mathbf{K}]^{-1}\mathbf{B}\mathbf{K}.$$

At every step  $n$ , write

$$(58) \quad \hat{\mathbf{u}}_n = \mathbf{u} + \mathbf{M}[\mathbf{u} - \widehat{\mathbf{u}}_n - \mathbf{1}].$$

Consider a norm  $\|\cdot\|$  on  $IR^{n_u}$  and consider  $\|\cdot\|$  the subordinate norm on a  $n_u \times n_u$  matrix such that for any  $\mathbf{u} \neq \mathbf{0}$ ,

$$\|\mathbf{M}\| = \sup_{\mathbf{u}} \frac{\|\mathbf{M}\mathbf{u}\|}{\|\mathbf{u}\|}.$$

Then,  $\|\mathbf{M}\mathbf{u}\| \leq \|\mathbf{M}\|\|\mathbf{u}\|$  and if  $\|\mathbf{M}\| < 1$ , then  $\hat{\mathbf{u}}_n$  converges to  $\mathbf{u}$ .

## 7. CONCLUSION

An asserted advantage of moment-based structural econometrics over reduced-form methods is that one can correctly identify policy-invariant parameters so that alternative policy options can be assessed. As we have shown, this approach, which generally treats policy changes as counterfactual zero probability exogenous events, violates rational expectations: Agents inside the structural model should understand that policy changes are positive probability endogenous events, which the econometric exercise is intended to inform. We examined the implications of this violation of rational expectations in moment-based econometric parameter inference, which serves a policy function. As shown, bias emerges unless the true value of the parameter justifies the status quo. If instead a policy change is justified, biased inference occurs. However, it was shown how rational expectations can be imposed in an internally consistent manner, yielding unbiased inference and optimal policy. Finally, it was shown that the switch from irrelevant estimation to policy-relevant estimation changes the nature of moment selection, with the focus switching from partial derivatives of moments to total derivatives, accounting for policy feedback.

The more general point illustrated by our analysis is that econometric methods should vary according to whether the estimation is passive or active in the sense of influencing policy decisions. Although the specifics of the transmission mechanism will differ, the essential problem highlighted by this article is that with active estimation, future endogenous policy will be correlated with the causal parameters to be estimated. If agents have rational expectations, this channel will bias structural inference if the inference-policy feedback effect is not taken into account. A potentially important direction for future research is to incorporate the policy control channel into the econometric tool kit.

## REFERENCES

- ADDA, J., AND R. COOPER, *Dynamic Economics* (Cambridge: MIT Press, 2003).
- BALLARD, C. L., J. B. SHOVEN, AND J. WHALLEY, "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States," *American Economic Review* 75 (1985), 128–38.
- BIANCHI, F., AND C. ILUT, "Monetary/Fiscal Policy Mix and Agents' Beliefs," *Review of Economic Dynamics* 26 (2017), 113–39.
- BLANCHARD, O., AND P. PORTUGAL, "What Hides Behind an Unemployment Rate: Comparing Portuguese and US Labor Markets," *American Economic Review* 91 (2001), 187–207.
- BLUME, L., D. EASLEY, AND M. O'HARA, "Characterization of Optimal Plans in Stochastic Dynamic Programs," *Journal of Economic Theory* 28 (1982), 221–34.
- BOUCHAUD, J.-P., P. KRUEGER, A. LANDIER, AND D. THESMAR, "Sticky Expectations and the Profitability Anomaly," *Journal of Finance* 74 (2019), 639–74.
- CHEMLA, G., AND C. A. HENNESSY, "Rational Expectations and the Paradox of Policy-Relevant Natural Experiments," *Journal of Monetary Economics* 114 (2020), 368–81.
- EJARQUE, J., AND P. PORTUGAL, "Labor Adjustment Costs in a Panel of Establishments: A Structural Approach," Discussion Paper 3091, IZA Bonn, 2007.
- ENTHOVEN, A. C., AND K. J. ARROW, "A Theorem on Expectations and the Stability of Equilibrium," *Econometrica: Journal of the Econometric Society* 24 (1956), 288–93.
- FARMER, R. E. A., D. F. WAGGONER, AND T. ZHA, "Understanding Markov-Switching Rational Expectations Models," *Journal of Economic Theory* 144 (2009), 1849–67.
- GALLANT, A. R., AND G. TAUCHEN, "Which Moments to Match?," *Econometric Theory* 12 (1996), 657–81.
- GOLDSTEIN, R., N. JU, AND H. LELAND, "An EBIT-Based Model of Dynamic Capital Structure," *Journal of Business* 74 (2001), 483–512.
- GOMES, J., "Financing Investment," *American Economic Review* 91 (2001), 1263–85.
- GORDON, R. H., AND J. K. MACKIE-MASON, "Tax Distortions to the Choice of Organizational Form," *Journal of Public Economics* 55 (1994), 279–306.
- GRUBER, J., AND J. RAUH, "How Elastic Is the Corporate Income Tax Base?," in Alan J. Auerbach, ed., *Taxing Corporate Income in the 21st Century* (Cambridge: Cambridge University Press, 2007), 140–63.
- HAMMERMESH, D. S., "Labor Demand and the Structure of Adjustment Costs," *American Economic Review* (1989), 674–89.
- HENNESSY, C. A., AND T. WHITED, "Debt Dynamics," *Journal of Finance* 60 (2005), 1129–65.
- HUBBARD, R. G., "Corporate Tax Integration: A View from the Treasury Department," *Journal of Economic Perspectives* 7 (1993), 115–32.
- KEANE, M., AND K. WOLPIN, "Introduction to the JBES Special Issue on Structural Estimation in Applied Microeconomics," *Journal of Business and Economic Statistics* 15 (1997), 111–14.
- , AND ———, "Estimating Welfare Effects with Forward Looking Agents," *Journal of Human Resources* (2002), 570–99.
- KYDLAND, F., AND E. PRESCOTT, "The Computational Experiment: An Econometric Tool," *Journal of Economic Perspectives* 10 (1996), 69–85.
- LELAND, H., "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance* 49 (1994), 1213–52.
- LUCAS, R., "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1 (1976), 19–46.
- MILLER, M. H., "Debt and Taxes," *Journal of Finance* 32 (1977), 261–75.
- MOYEN, N., "Investment-Cash Flow Sensitivities: Constrained vs Unconstrained Firms," *Journal of Finance* 59 (2005), 2061–92.
- NADEAU, S., "A Model to Measure the Effect of Taxes on the Real and Financial Decisions of the Firm," *National Tax Journal* 41 (1988), 467–81.
- ROBERDS, W., "Models of Policy under Stochastic Replanning," *International Economic Review* (1987), 731–55.
- RUST, J., "The Limits of Inference with Theory: A Review of Wolpin (2013)," *Journal of Economic Literature* 52 (2014), 820–50.
- SARGENT, T. J., "Autoregressions, Expectations, and Advice," *American Economic Review* 74 (1984), 408–15.
- , "Rational Expectations," in Steven N. Durlauf and Lawrence E. Blume, eds., *Macroeconomics and Time Series Analysis* (London: Palgrave Macmillan, 2010), 193–201.
- , "Rational Expectations," in J. Eatwell, M. Milgate and P. Newman, eds., *The New Palgrave Dictionary* (London: MacMillan Press, 1987), 76–85.
- SARGENT, T., *The Conquest of American Inflation* (New Jersey: Princeton University Press, 2018).
- SARGENT, T. J., AND N. WALLACE, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics* 2 (1976), 169–83.

- SIMS, C. A., "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity* 1 (1982), 107–64.
- , AND T. ZHA, "Were There Regime Switches in US Monetary Policy?," *American Economic Review* 96 (2006), 54–81.
- STIGLITZ, J. E., "Taxation, Corporate Financial Policy, and the Cost of Capital," *Journal of Public Economics* 2 (1973), 1–34.
- WOLPIN, K., *The Limits of Inference Without Theory* (Cambridge: MIT Press, 2013).