

BEYOND RANDOM ASSIGNMENT:
CREDIBLE INFERENCE AND EXTRAPOLATION
IN DYNAMIC ECONOMIES*

Christopher A. Hennessy

London Business School, CEPR, and ECGI

Ilya A. Strebulaev

Stanford GSB and NBER

August 2018

Abstract

We derive analytical relationships between shock responses and theory-implied causal effects (comparative statics) in dynamic settings with linear profits and linear-quadratic stock accumulation costs. For permanent profitability shocks, responses can have incorrect signs, undershoot, or overshoot depending on whether the realized changes are small, large, or have opposite sign relative to expected changes. For profitability shocks that are i.i.d., uniformly distributed, binary, or unanticipated and temporary, there is attenuation bias, which exceeds 50% under plausible parameterizations. We derive a novel sufficient condition for profitability shock responses to equal causal effects: martingale profitability. We establish a battery of sufficient conditions for correct sign estimation, including stochastic monotonicity. Simple extrapolation/error correction formulae are presented.

*We thank Michael Roberts and Philip Bond (the editors), and two anonymous referees for valuable input. We also thank Manuel Adelino, Antoinette Schoar, and Luke Taylor (discussants), as well as seminar participants at Stanford, Wharton, LBS, Duke, Boston University, LSE, UNC, UBC, Maryland, NC State, Imperial College, Simon Fraser, Koc, INSEAD, VGSF, SFI, SFS Cavalcade, the Causal Inference in Finance and Economics Conference, and the Stanford Conference on Causality in the Social Sciences. Disclosures: Hennessy received funding from the European Research Council related to this project. Strebulaev received no funding related to this project.

In applied microeconomics fields such as corporate finance, *theory-implied causal effects* are generally derived using comparative statics. Heckman (2000) writes, “Comparative statics exercises formalize Marshall’s notion of a *ceteris paribus* change which is what economists mean by a causal effect.” Athey, Milgrom and Roberts (1998) write, “most of the testable implications of economic theory are comparative static predictions.” In assessing implications of tradeoff theory, Welch (2010) writes, “The comparative statics analysis are simple and obvious: The firm should have more debt if the interest rate is higher, if the corporate tax rate is higher, and the cost of distress is lower.”

The objective of much econometric work is to empirically estimate signs and magnitudes of causal effects. The estimates can be compared with theory-predicted causal effects and can also inform welfare analysis. Theories incorrectly predicting signs are informally labeled “falsified” and those with small estimated causal effects are informally labeled “second-order.” Small empirical estimates of causal effects have been used to argue that government-induced behavior distortions generate small deadweight losses (see e.g. Slemrod (1990) and Aaron (1990)).

In their textbook, *Mostly Harmless Econometrics: An Empiricist’s Companion* (MHE below), Angrist and Pischke (2009) imply that the sole/primary barrier to correct empirical estimation of causal effects is avoiding endogeneity and selection bias: “The goal of most empirical research is to overcome selection bias, and therefore to have something to say about the causal effect of a variable.” They also state, “A principle that guides our discussion is that most of the estimators in common use have a simple interpretation that is not heavily model dependent.” The perceived conjunction of credibility and simplicity of interpretation has led to a surge in the popularity of the MHE methodology.

A limitation of the MHE methodology in practice is that it is silent on how to deal with the inherent complications arising from dynamic uncertainty. In particular, while there is much discussion about shock exogeneity, there is rarely discussion, let alone formal estimation, of the data generating process responsible for the exploited shocks. Moreover, there is rarely any formal attempt to: map observed shock responses back to theory-implied causal effects (theory testing); extrapolate shock responses across stochastic environments (external validity); or estimate causal parameters entering welfare calculations. In some cases, researchers may have a prior that dynamic uncertainty is a second-order concern. In still other cases, researchers are aware that inference is somehow clouded, but do not know the exact nature of the problems nor how to address them at

an operational level without resorting to full-scale structural estimation.

Issues arising from dynamic uncertainty are avoided if one claims an exploited shock is unanticipated and permanent. However, this common assumption has a host of problems. First, confining attention to truly rare events means that timely evidence is unlikely to exist. Second, the notion that an exploited shock is permanent generally contradicts the underlying motivation for empirical work: to help optimize policy in the future and/or to forecast responses to future shocks. Third, inspection of time-series for many policy variables such as the effective corporate tax rate and the real minimum wage reveals changes to be the norm, not the exception. Finally, there is a fundamental difference between shocks never occurring versus shocks taking place infrequently, say, once a decade in expectation. In the latter case, the arrival of a shock today might be casually treated as a “surprise.” Nevertheless, as we show, the bias arising from expectations in such cases can still be quite large.

What special challenges does dynamic uncertainty pose for researchers performing shock-based inference, how important are they, and what can be done to overcome them without resorting to full-scale structural estimation? In order to address these questions, we develop a canonical model of firm decision-making flexible enough to capture a broad range of commonly-studied dependent variables and alternative causal theories. In our baseline model, a firm dynamically accumulates a (divisible) stock, such as capital, inventories, patents, debt, cash, or employees, which provides a (linear) flow of benefits over time while facing (linear-quadratic) accumulation costs.

Using the model as our laboratory, we consider an econometrician whose objective is to infer whether the shock responses she observes are consistent with the theory-implied causal effects. Fortunately, we know the theory-implied causal effects, which we compute using comparative statics. In contrast, our econometrician is forced to rely upon shock responses for inference. But critically, we consider a setting in which the endogeneity challenges to which MHE devotes exclusive attention are a non-issue. Specifically, we assume shocks take the form of exogenous Markov chain transitions. Examining the relationship between theory-implied causal effects and shock responses in this laboratory allows us to assess whether the MHE tool-kit suffices for inference of theory-implied causal effects in dynamic settings.

Our primary focus is on inference when the econometrician exploits shocks that bring about changes in the flow of periodic benefits generated by the stock, e.g. shocks to taxes, regulations,

competition, productivity, input prices, or yields on cash/debt. To begin, here we derive a set of tractable analytical formulae relating shock responses to theory-implied causal effects. Each formula can be used in performing back-of-the-envelope bias calculations, or inverted to correct for bias.

The bias formulae reveal troubling implications. For example, if the shock is, say binary, then shock responses can represent severely downward biased estimates of theory-implied causal effects. For example, under our baseline parameterization, attenuation bias exceeds one-half the causal effect if the shock probability exceeds just 5.26%. Worse still, we show shock responses can be biased upwards and even have signs opposite to that of the causal effect. For example, if a shock is permanent, the shock response will overshoot the causal effect if the realized change is opposite in sign to the expected change. Intuitively, objectively good (bad) news becomes great (terrible) news if bad (good) news was expected. Worse still, the shock response will have the wrong sign if the realized change is sufficiently small relative to the expected change. Intuitively, objectively good (bad) news becomes bad (good) news if even better (worse) news was expected.

The problems these examples highlight are not theoretical curiosities. To illustrate, using the time-series of historical effective tax rates, we estimate a Markov chain approximating the data generating process confronting corporations. We then analytically compare simulated shock responses with theory-implied causal effects. Under our baseline parameterization, we find that attenuation bias is severe, with shock responses generally falling just below one-half the theory-implied causal effect. Intuitively, real-world transience in tax rates makes firms much less responsive to shocks than would be the case if the shocks were unanticipated and permanent. Such severe magnitude bias has important policy implications. After all, as we show, the same 50% attenuation bias exhibited by shock responses carries over directly to estimated tax-elasticity and excess burden calculations.

To make sense of shock responses, mapping them back to theory-implied causal effects or extrapolating them, one must move beyond exogeneity by estimating the stochastic process for the shocks (the “data generating process” below). We show analytically how, with this estimation accomplished, theory-implied causal effects can be recovered from shock responses, so that empirical evidence can speak directly to theoretical predictions or fed directly into welfare calculations. We also show analytically how shock responses can be extrapolated within and across data generating processes overcoming the challenge that dynamic uncertainty poses to external validity.

In some circumstances, the econometrician will be unable to estimate the underlying data generating process, but will want to argue her setting satisfies some identifying assumption that guarantees equality of shock responses and causal effects. As we show, one valid identifying assumption is that all state-contingent shock probabilities are infinitesimal, i.e. shocks are almost perfectly unanticipated and permanent. However, we also show that under the classical linear-quadratic technology, another valid identifying assumption is that the flow of benefits provided by the stock be a martingale. In some applied settings, this may be viewed as a more plausible identifying assumption since a stochastic process can change frequently and still be a martingale.

In some cases, an empiricist will set a lower bar and aim to at least avoid incorrect estimation of causal effect signs. Here we derive a battery of sign identification assumptions that can be invoked: i.i.d. shocks; unanticipated temporary shocks; binary shocks; transitory uniformly distributed shocks; a permanent change opposite in sign to the expected change; a permanent change equal in sign but with absolute value exceeding the expected change; or stochastic monotonicity. Indeed, *stochastic monotonicity* may be a reasonable sign identification assumption to invoke in many applied settings: An increase in the causal variable today implies a first-order stochastic dominant shift in the variable at all future dates.

We also examine shocks to the stock variable price and shocks to adjustment costs. Here we find that under the linear-quadratic technology measured shock responses are equal to theory-implied causal effects regardless of the underlying stochastic process. However, this finding is due to a special property of the linear-quadratic technology: the amount of stock accumulated this period has no effect on the feasibility, cost, and benefit flow arising from accumulation in future periods.

To illustrate, as extensions we consider the relationship between shock responses and causal effects under alternative technologies. As discussed above, with linear benefits and quadratic adjustment costs, the stochastic process matters for interpreting benefit shocks but not for interpreting shocks to the price of the stock variable. We show that if the firm instead faces concave benefits to the stock variable with zero adjustment cost, these conclusions are reversed. In particular, in this case, the stochastic process for the shocks is irrelevant for the interpreting benefit shocks but critical for interpreting price shocks. Thus, it is apparent from our analysis that correctly interpreting shock responses requires making joint assumptions about underlying stochastic processes and real technologies, in addition to the traditional orthogonality assumption.

We turn now to related literature. Bowen, Frésard, and Taillard (2016) find that in the top-three finance journals, the share of empirical corporate finance papers using what they term “identification technologies” has risen from roughly 0 percent in the late 1980’s to over 50 percent by 2012, with 126 published papers using difference-in-difference techniques applied to quasi-natural experiments.

The most common identification technologies are “shock-based,” with inferences being drawn from responses to changes in causal variables. For example, an extensive literature studies capital structure responses to tax code changes.¹ Shock-based inference is also pervasive in corporate governance research. In an exhaustive analysis, Atanasov and Black (2016) document that shock-based designs have risen from 4% of empirical governance papers in 2001-2006 to 11% in 2007-2011. As surveyed by Karpoff and Wittry (2015), changes in state-level anti-takeover laws are often exploited as a source of quasi-exogenous variation in governance pressure.² As a final example, shock-based inference is common in studying real investment which, along with leverage, is perhaps the most commonly-studied dependent variable in empirical corporate finance.³

Our analysis is in the spirit of Lucas (1976) who points out the perils of extrapolating macroeconomic evidence across policy rules. Our focus is on the inadequacy of the random assignment assumption taken in isolation. Naturally, Lucas does not analyze the relationship between treatment responses under dynamic random assignment and theory-implied causal effects, nor derive sufficient conditions for correct identification of causal effect signs and magnitudes, nor derive formulae for error correction and extrapolation.

Abel (1982) analyzes effects of permanent versus temporary tax policies with no policy variable uncertainty. Specifically, he assumes, in potential violation of rational expectations, that the policy change is completely unanticipated, with policy being binary and deterministic after the complete surprise policy change. Incorporating rational expectations, as well as the probability distribution of shocks, is essential for our results. House and Shapiro (2008) also ignore uncertainty and present response approximations assuming arbitrarily short-lived policy changes. In their setting,

¹See Gordon and MacKie-Mason (1990); Givoly et al. (1992); Gordon and MacKie-Mason (1997); Chetty and Saez (2004); Hanlon and Joopes (2014); Heider and Ljungqvist (2015); Schepens (2016); and Schandlbauer (2017).

²See Gordon and Hanka (1999).

³See Cummins et al. (1994); Yagan (2015); Bertrand and Mullainathan (1999, 2003); Atanasov (2013); Greenstone (2002); Alesina et al. (2005); Kilian (2008); Gan (2007); Chaney et al. (2012); Banerjee et al. (2015); Crepon et al. (2015); and Mukherjee et al. (2017).

the shadow value of capital is left unchanged by policy changes. In contrast, we analyze the behavior of shadow values, and optimal accumulation, allowing for policy uncertainty and expected regime lives of arbitrary duration.

Auerbach (1986) and Auerbach and Hines (1988) present investment Euler equations under stochastic tax rates. Cummins, Hassett and Hubbard (1994) estimate adjustment cost parameters based on tax experiments. They assume each tax change is a complete surprise and is viewed as being permanent, in violation of rational expectations. Hassett and Metcalf (1999) present a real options model with a one-time investment which they use to assess whether uncertainty regarding tax credits encourages investment. Gourio and Miao (2008) numerically compare effects of permanent and temporary dividend tax cuts.

The closest paper to our own is that by Keane and Wolpin (2002). One of the underlying messages of the papers is similar, with both stressing the necessity of accounting for dynamics and uncertainty in statistical inference. However, there are numerous important differences. First, in terms of context, they analyze a granular dynamic model of contraceptive use and welfare participation. We instead offer a more general analysis of the effect of dynamic uncertainty on shadow values, the key determinant of optimal accumulation of a broad array of commonly-studied stock variables. Second, they offer numerical solutions featuring polynomial approximations while we present analytic solutions amenable to direct analysis. Third, and related to this point, we present simple analytic expressions for biases, and bias probabilities, stated in terms of expectations and shock probabilities. Finally, we present necessary and sufficient conditions for correct signs and/or magnitudes. They do not address these questions.

Our work is in the spirit of a paper by Chetty (2012) who also attempts to straddle reduced-form and structural work. He writes, “The identification of structural parameters of stylized models is one of the central tasks of applied economics.” Chetty analyzes how to recover structural elasticity parameters if there are transaction costs or inattention. We instead develop a model to understand and correct empirical estimates derived from random assignment in dynamic settings, with recovery of adjustment cost parameters being an interim step.

Our paper is related to but distinct from the literature contesting the external validity of estimates derived from instrumental variables, e.g. Heckman (1997) and Deaton (2010). To begin, in the setting we consider, there is absolutely no instrumentation and by construction no possibility

of self-selection, so the oft-discussed issues surrounding the correct interpretation of LATE are extraneous. Second, issues concerning heterogeneous treatment effects are also extraneous to our analysis.

The remainder of the paper is organized as follows. Section I describes the economic setting. Section II derives the relationship between shock responses and theory-implied causal effects. Section III presents identifying assumptions. Section IV shows how shock responses can be extrapolated. Section V presents a numerical example. Section VI considers alternatives to the linear-quadratic technology. Section VII concludes.

I. Theory and Models

Our objective is to develop an analytical framework allowing researchers to both understand and bridge the gap between underlying theories and empirical tests of those theories relying upon exogenous time-series variation in variables for the purpose of causal identification.

We begin by writing down a classical linear-quadratic model of how a firm will optimally accumulate a stock variable over time. As demonstrated below, this model will be general and flexible enough to encompass a broad set of causal theories and dependent variables. Once the model is in-hand, we perform comparative statics with respect to its causal parameters.⁴ In this way, we mimic the methodology theorists and empiricists generally rely upon when attempting to formally flush out the testable implications of underlying theories. We then develop a companion model that captures the very same underlying theoretical tradeoffs as the original model, but that mimics the real-world data generating process exploited by the econometrician.

The body of the paper presents only those equations necessary for the empirical implications. Formal solutions, including derivation of the value function, are relegated to the appendix.

A. Theory-Implied Causal Effects

This subsection presents a general linear-quadratic model of optimal accumulation of a stock variable. The stock variable provides a linear benefit flow, and the firm faces linear-quadratic costs of adjusting the stock.

⁴As in Heckman (2000), a “causal parameter” is a model parameter whose effect on outcomes we would like to predict. Causal parameters are a subset of a model’s structural parameters.

In the interest of generality, we treat the model’s benefit and cost parameters as primitives. Depending on the theory being considered, alternative factors would cause changes in these parameters. For example, in tax-based theories of investment and financial structure, a higher corporate tax rate is associated with a lower flow of benefits from the accumulation of real and financial assets. Or to take another example, in theories emphasizing transactions costs, streamlining the bond flotation process through shelf-registration would be associated with lower costs to adjusting debt stocks. Similarly, labor market deregulation would be associated with lower costs to adjusting the number of employees.

Time is discrete and horizons are infinite. Agents are risk-neutral and discount at rate $r > 0$. The firm accumulates a stock s that decays at rate $\delta \geq 0$. The firm starts the current period with stock s and receives an end-of-period net after-tax benefit inflow equal to bs' where s' is the end-of-period stock and $b > 0$. Notice, corporate saving and borrowing are subsumed here with $s > 0$ corresponding to a positive liquid asset balance and $s < 0$ corresponding to debt.

The firm can buy and sell the stock variable at price p . The variable a denotes *accumulation*. The law of motion for the stock variable is:

$$s' = (1 - \delta)s + a. \tag{1}$$

We do not impose a non-negativity constraint on the stock. Indeed, in order for the model to accommodate corporate borrowing, we must admit $s < 0$. A non-negative steady-state stock can be guaranteed by imposing $p \leq b/(r + \delta)$. Alternatively, one can simply assume the initial stock held by the firm is arbitrarily large.

The stock variable price p represents the second causal parameter in the model. Finally, there are quadratic costs to adjusting the stock equal to $a^2/2\xi$. Here $\xi > 0$ is the third and final causal parameter in the model capturing the ease of adjusting the stock. Importantly, this linear-quadratic framework can be understood as an implicit formal foundation for much empirical work since, as shown below, it maps into the classical linear causal effects econometric model. A more general class of cost functions will be considered as an extension in Section 7.

A wide range of dependent variables is subsumed in this model. The most obvious interpretation of the stock variable is that it represents tangible or intangible capital, e.g. PPE, patents, brand value, or reputation. Alternatively, s can represent a buffer stock of net working capital, such

as cash or inventories that provide convenience yields. One can think of a negative value of s as representing default-free debt. As a final example, following Hammermesh (1989), s can represent the number of employees.

Despite its simplicity, a broad range of real-world causal effects can be captured in this model. Consider first causal factors that change the benefit inflow b . Changes in taxes and regulations change profits per unit of capital. Demand and supply shocks change flexible input prices, affecting operating profits.⁵ Consider next just some of the causal factors that can change the stock price p , as well as the ease of stock adjustment ξ . Tariffs, regulation and competition change equilibrium prices of capital goods. Government affects the net price of capital through investment tax credits and R&D tax credits. The imposition of a zoning restriction changes the price of real estate, as do natural disasters. Labor market regulations affect the ease (ξ) of hiring and firing employees, and securities market regulations affect the ease of adjusting debt stocks.

At each point in time, it is optimal for the firm to increase accumulation up to the point that the marginal value of the stock variable, call it q , is just equal to the marginal cost. Equating marginal benefits and costs we have:

$$q = p + \xi^{-1}a^{**} \Rightarrow a^{**} = \xi(q - p). \quad (2)$$

Notice, in the preceding equation double-stars have been used to denote the optimal policy obtained in the present setting which will contemplate changes in the causal parameter vector (b, p, ξ) by way of standard comparative statics analysis.

Accounting for the fact that the stock variable decays at rate δ , the Gordon growth formula can be used to pin down the present value of the flow of benefits. Anticipating, the next subsection will consider b to be a Markov process with i indexing the state. With this in mind we write

$$b = b_i \Rightarrow q = \frac{b_i}{r + \delta} \equiv q_{\infty}^i. \quad (3)$$

For the remainder of the paper the variable q_{∞}^i will be labeled *permanent* q . It measures the shadow value of a unit of the stock variable if the benefit inflow variable is permanently fixed at b_i .

The two preceding equations imply the optimal accumulation policy is:

$$a^{**}(b, p, \xi) = \xi \left[\frac{b}{r + \delta} - p \right]. \quad (4)$$

⁵See Abel and Eberly (1994) for calculations.

We then obtain the following theory-implied causal effects:

$$\begin{aligned}\frac{\partial a^{**}(b, p, \xi)}{\partial p} &= -\xi \\ \frac{\partial a^{**}(b, p, \xi)}{\partial \xi} &= \frac{b}{r + \delta} - p \\ \frac{\partial a^{**}(b, p, \xi)}{\partial b} &= \frac{\xi}{r + \delta}.\end{aligned}\tag{5}$$

Intuitively, the theory here predicts that optimal stock accumulation is: decreasing in the price p of the stock; increasing (decreasing) in the efficiency of stock adjustments when it is optimal to increase (decrease) the stock; and increasing in the periodic benefit flow b .

In the case of discrete changes, as opposed to infinitesimal changes, causal effects are evaluated by computing the difference between optimal policies evaluated at different values of the causal parameters. We have the following definition of a discrete causal effect associated with a move from causal parameter vector i to causal parameter vector j :

$$CE_{ij} \equiv a^{**}(b_j, p_j, \xi_j) - a^{**}(b_i, p_i, \xi_i).\tag{6}$$

Substituting the optimal policies from equation (4) into the preceding equation we obtain the following theory-implied causal effects:

$$\begin{aligned}p &: CE_{ij}^p = -\xi(p_j - p_i) \\ \xi &: CE_{ij}^\xi = \left(\frac{b}{r + \delta} - p\right)(\xi_j - \xi_i). \\ b &: CE_{ij}^b = \xi \left(\frac{b_j - b_i}{r + \delta}\right).\end{aligned}\tag{7}$$

Alternatively, following Heckman (2000)), the fact that the causal effects are linear (equation (5)) implies that the preceding expressions for discrete causal effects can be computed by multiplying partial derivatives by the contemplated discrete change in causal parameters.

The steady state stock here, call it s^{**} , is such that optimal periodic accumulation is just sufficient to cover depreciation on the stock. Thus:

$$s^{**} = \frac{a^{**}}{\delta} \Rightarrow \Delta s^{**} = \frac{\Delta a^{**}}{\delta}.\tag{8}$$

Even though there are no shocks here, there will be dynamics. Specifically, if the stock is initially below (above) s^{**} , the stock will increase (decrease) since accumulation is greater than (less than)

depreciation. Finally, notice that the comparative static properties of the flow variable a^{**} carry over to the stock variable s^{**} , being scaled by δ^{-1} . Thus, under the linear-quadratic technology, causal effects are linear whether they are expressed in terms of the flow variable (a^{**}) or the steady-state stock value (s^{**}).

B. Shocks and Firm Decisions

We consider now a companion economy where the underlying theoretical tradeoff is exactly the same as in the preceding subsection: The optimal stock accumulation by the firm weighs linear-quadratic accumulation costs against the periodic benefit inflow b provided by the stock. However, we construct this companion economy to mimic the data generating process actually being exploited by the econometrician who performs shock-based inference. One can think of the econometrician as observing firms before and after an exogenous shock, or as comparing shocked firms with otherwise identical non-shocked firms.

By construction, we strip our laboratory economy of selection and endogeneity bias. In particular, the evolution of the three causal variables b , p and ξ is governed by an exogenous N -state Markov chain. The shocks are also clean in the sense that only one of the three causal variables is changing over time, with the other two variables assumed to be constant. This ensures that inferences regarding firm responsiveness to a specific causal variable cannot be clouded by changes in some other causal variable.

Let state 1 denote the best state in the sense of being associated with the highest benefit inflow, highest stock adjustment efficiency, and lowest stock variable price:

$$\begin{aligned} b_1 &\geq \dots \geq b_N & (9) \\ \xi_1 &\geq \dots \geq \xi_N \\ p_1 &\leq \dots \leq p_N. \end{aligned}$$

When considering experiments involving shocks to a given causal variable, say b , there will be N strict inequalities for that variable, e.g. $b_1 > \dots > b_N$, and N equalities for the remaining two variables which are then assumed to be constant over time.

The timing convention is as follows. If the current (beginning of period) state is i , then the firm makes its current accumulation decision knowing that: the price of the stock is p_i , the ease

of stock adjustment is ξ_i , and the end-of-period benefit inflow is $b_i s'$ where s' reflects the current period's accumulation decision in accordance with equation (1). Further, if the current state is i , then with probability Λ_i there will be a shock causing the state to be different at the start of the next period. Finally, the conditional probability of a transition to state j , given that a transitional shock out of state i has taken place, is denoted Π_{ij} , with $\Pi_{ii} = 0$.

Since the underlying theory here has been assumed to be the same as in the preceding subsection, the optimal accumulation policy again equates the marginal value of the stock variable with the marginal cost of accumulation. However, we must now depart from equation (2) and account for state-contingency in the marginal stock value, stock price, and efficiency of stock adjustment. If the current state is i , then the optimal state-contingent accumulation policy, call it a_i^* , solves:

$$q_i = p_i + \xi_i^{-1} a_i^* \Rightarrow a_i^* = \xi_i (q_i - p_i). \quad (10)$$

Notice, in the preceding equation, single stars are used to denote optimal policy for a firm occupying the same economy as the econometrician, an economy featuring exogenous shocks to causal variables. Further, for the remainder of the paper, the variable q_i will be labeled *transient q* since it measures the value of a unit of the stock variable in an economy in which the benefit b is subject to shocks.

Consider now the marginal value of a unit of the stock variable if the current state is i . At the end of the period, one unit of accumulation will generate a known benefit inflow b_i . In addition, $1 - \delta$ units of decayed stock will carry over to the subsequent period. However, the state in the subsequent period is uncertain. With probability $(1 - \Lambda_i)$ the state will not change and so each unit of the decayed stock will have value q_i . With probability $\Lambda_i \Pi_{ij}$ the state will change to j and each unit of decayed stock will have value q_j . It follows that state-by-state, the marginal value of the stock satisfies:

$$\begin{aligned} q_1 &= \left(\frac{1}{1+r} \right) \left[b_1 + (1-\delta) \left((1-\Lambda_1)q_1 + \Lambda_1 \sum_{j \neq 1} \Pi_{1j} q_j \right) \right] \\ &\dots \\ q_N &= \left(\frac{1}{1+r} \right) \left[b_N + (1-\delta) \left((1-\Lambda_N)q_N + \Lambda_N \sum_{j \neq N} \Pi_{Nj} q_j \right) \right]. \end{aligned} \quad (11)$$

Equation (11) is a simple linear system. Solving it, we find that the marginal value of the stock

state-by-state is:

$$\begin{bmatrix} q_1 \\ \dots \\ q_N \end{bmatrix} = [\mathbf{T}^{-1}] \begin{bmatrix} b_1 \\ \dots \\ b_N \end{bmatrix} \quad (12)$$

where \mathbf{T} denotes the following *augmented transition matrix*:

$$\mathbf{T} \equiv \begin{bmatrix} r + \delta + (1 - \delta)\Lambda_1 & -(1 - \delta)\Lambda_1\Pi_{12} & \dots & -(1 - \delta)\Lambda_1\Pi_{1N} \\ -(1 - \delta)\Lambda_2\Pi_{21} & r + \delta + (1 - \delta)\Lambda_2 & \dots & -(1 - \delta)\Lambda_2\Pi_{2N} \\ \dots & \dots & \dots & \dots \\ -(1 - \delta)\Lambda_N\Pi_{N1} & -(1 - \delta)\Lambda_N\Pi_{N2} & \dots & r + \delta + (1 - \delta)\Lambda_N \end{bmatrix}. \quad (13)$$

Conveniently, the preceding linear system accounts for all shocks over time and also performs the necessary discounting.

In the special case in which the benefit flow only has two possible values, equation (12) can be written as:

$$q_i = \frac{b_i}{r + \delta} + \frac{(1 - \delta)\Lambda_i(b_j - b_i)}{(r + \delta)[r + \delta + (1 - \delta)(\Lambda_i + \Lambda_j)]}. \quad (14)$$

II. Shock Responses and Causal Effects

This section describes the relationship between shock responses and causal effects. Conveniently, in contrast to the real-world environment confronting the econometrician, we here know the theory-implied causal effects (equations (5) and(7)). Thus, we will be able to determine the conditions under which shock responses recover the sign and/or magnitude of the causal effects. Moreover, when there is indeed a wedge, we will be able to derive expressions allowing econometricians to back out the causal effect implied by a measured shock response.

A. Defining Shock Responses

As discussed above, the econometrician occupies a stochastic economy where the firm's optimal state-contingent accumulation policy (a_i^*) is described by equation (10). The econometrician then relies on observed shock responses for causal inference. The shock response associated with a transition from state i to j is denoted SR_{ij} . We have:

$$SR_{ij} \equiv a_j^* - a_i^*. \quad (15)$$

Substituting the optimal policies from equation (10) into the preceding equation, it follows that the econometrician in the stochastic economy will observe the following respective shock responses:

$$\begin{aligned}
p & : SR_{ij}^p = -\xi(p_j - p_i) \\
\xi & : SR_{ij}^\xi = (q - p)(\xi_j - \xi_i) = \left(\frac{b}{r + \delta} - p\right)(\xi_j - \xi_i). \\
b & : SR_{ij}^b = \xi(q_j - q_i).
\end{aligned} \tag{16}$$

B. Inference: Shocks to Stock Variable Price and Adjustment Costs

It follows directly from equation (16)) that the shock responses observed by the econometrician are identical to the theory-implied causal effects (equation (7)) associated with changes in the causal parameters related to the stock price and efficiency of stock adjustment. We thus have the following result.

Proposition 1 *With exogenous changes in either the price p of the stock, or exogenous changes in stock adjustment costs (ξ), measured shock responses directly recover theory-implied causal effects regardless of the underlying stochastic process.*

To understand the preceding result, some points are worth noting. First, since benefit inflows are, by assumption, linear in the stock variable, the amount of accumulation the firm chooses this period has no effect on the benefits provided by units of stock added in future periods. Second, in contrast to a real-options setting, in a neoclassical setup such as ours, stock accumulation this period has no effect on the feasibility of stock accumulation in future periods. Finally, since accumulation costs are, by assumption, linear-quadratic in a , the amount of accumulation the firm chooses this period has no effect on the cost of accumulation in future periods.

In such an environment, the firm can behave myopically when it evaluates paying the stock variable price or incurring adjustment costs. Thus, the law of motion for these variables becomes irrelevant to the firm's stock accumulation decision in the current period. By way of contrast, in Section VI it is shown that the law of motion does indeed become relevant for the interpretation of price shock responses if one instead considers a setting with concave benefits and zero adjustment costs. Thus, the preceding proposition is best viewed as a potentially useful identifying assumption, yet one that relies on strong functional form assumptions.

C. Inference: Shocks to Intertemporal Benefit Flows

This subsection considers econometric inference in economies where there are shocks to the benefit variable b .

To begin the formal analysis, we note that theory-implied causal effects (equation (7)) can be recovered from observed shock responses provided one has estimated the law of motion for the shocks. In particular, the following proposition provides the general mapping formula.

Proposition 2 *With exogenous shocks to the intertemporal benefit flow b , the ratio of causal effect to shock response is given by*

$$\frac{CE_{ij}^b}{SR_{ij}^b} = \frac{\xi(\Delta q_\infty)}{\xi(\Delta q)} = \frac{(b_j - b_i)/(r + \delta)}{q_j - q_i} \quad (17)$$

where q_i and q_j are determined by the law of motion for shocks per equation (12).

The preceding proposition describes a general algorithm for recovering causal effects from measured shock responses. Intuitively, in order to infer causal effects one must scale the observed shock response by the ratio of the change in permanent q_∞ (equation (3)) to the change in transient q (equation (12)). The key message delivered by the proposition is simple: Articulating/estimating the underlying stochastic process describing shocks is a necessary precursor for extracting their economic meaning. Section V provides an example of how to make Proposition 2 operational in applied settings.

As reflected in the preceding proposition, under the posited linear-quadratic technology, the adjustment cost parameter drops out of the bias ratio formula since both causal effects and shock responses are multiples of ξ . However, rearranging terms in the proposition, we see that bias, if measured as the difference between shock responses and causal effects increases linearly in ξ .⁶ In particular, rearranging terms in the preceding proposition, we find:

$$BIAS_{ij}^b \equiv SR_{ij}^b - CE_{ij}^b = \xi \left[(q_j - q_i) - \frac{b_j - b_i}{r + \delta} \right]. \quad (18)$$

Intuitively, greater ease of stock adjustment (ξ) will magnify bias associated with the wedge between transient and permanent q .

⁶Notice, transient q does not depend upon ξ under the posited linear-quadratic technology.

The preceding proposition also sheds light on structural econometric work by Cummins, Hassett and Hubbard (CHH below) (1994) who offer an estimate, call it $\widehat{\xi}$, of the ease of stock adjustment, which has true value ξ . If the shock generating process were known, equation (12) could be used to solve for the true shadow value of capital in each state. Exploiting equation (16), the true value of the ease of adjustment parameter could then be inferred from an observed shock response by solving:

$$\xi = \frac{SR_{ij}^b}{q_j - q_i}. \quad (19)$$

CHH impute shadow values of capital under the assumption that each tax change comes as a complete surprise and is viewed as being permanent. That is, the shadow value of capital in the preceding equation is treated as q_∞ as defined in equation (3). From equations (17) and (19) it follows that this imputation procedure delivers the parameter estimate

$$\widehat{\xi} = \frac{SR_{ij}^b}{q_\infty^j - q_\infty^i} = \frac{q_j - q_i}{\underbrace{(b_j - b_i)/(r + \delta)}_{SR/CE} \underbrace{q_j - q_i}_\xi} = \left(\frac{SR_{ij}^b}{CE_{ij}^b} \right) \times \xi. \quad (20)$$

At an operational level, the preceding equation illustrates how all the results we present on the ratio of shock responses to causal effects have a direct mapping to structural parameter inference. In particular, the CHH imputation procedure yields biased estimates of the parameter ξ , capturing ease of stock adjustment, precisely when shock responses yield biased estimates of theory-implied causal effects. To illustrate the practical implications, consider instances where reduced-form approaches, relying on shock responses, produce downward biased estimates of causal effects ($SR/CE \in (0, 1)$). In those same instances, structural approaches, using the CHH imputation method, would produce downward biased estimates of the ease of stock adjustment. Econometricians using the former approach erroneously conclude the underlying causal theory is “second-order” while structural econometricians will erroneously conclude stock adjustment will be “sluggish” due to extremely high adjustment costs.

In many instances, the econometrician will not have sufficient information to utilize Proposition 2, which requires knowledge or estimation of the stochastic process for the shocks. Nevertheless, the econometrician may have a rough sense of the data generating process. Our objective in the remainder of this section is to provide useful take-aways in such settings. To begin, suppose econometric inference is to take place in a setting in which the benefit variable b can assume one

of two possible values. For example, one might think of certain regulations as resulting in a binary flow of profits, e.g. a particular business practice or production input is either allowed or prohibited. Alternatively, one can think of mapping different genders and political parties into different values for pecuniary and non-pecuniary benefits. Substituting the binary shadow value formula from equation (14) into equation (20), we obtain the following proposition.

Proposition 3 *If the benefit variable is binary, with $b \in \{b_i, b_j\}$, the causal effect implied by an observed exogenous shock response is given by*

$$CE_{ij}^b = SR_{ij}^b \times \left[1 - \frac{(1 - \delta)(\Lambda_i + \Lambda_j)}{r + \delta + (1 - \delta)(\Lambda_i + \Lambda_j)} \right]^{-1},$$

with $sgn(CE_{ij}^b) = sgn(SR_{ij}^b)$ and $|CE_{ij}^b| > |SR_{ij}^b|$ if $\Lambda_i + \Lambda_j > 0$.

Notice, the preceding proposition shows that if b is a binary random variable, there is always bias, and this bias takes the form of attenuation. That is, in binary settings, shock responses necessarily have the same sign as the theory-implied causal effect but are smaller in absolute value. This result should be of comfort for those researchers who care primarily about whether the sign of a shock response is consistent with the underlying theory-implied comparative statics.

On the other hand, if the goal is to estimate economic magnitudes, as is necessary for policy analysis, or if the goal is assessing whether particular theories are “first-order,” then Proposition 3 raises important concerns. After all, the proposition implies that attenuation bias will be quite severe even in settings with low shock probabilities. As an example, suppose the econometrician inhabits an economy with $r = 2.5\%$ and $\delta = 7.25\%$, our baseline parameter assumptions. This real interest rate assumption follows Hennessy and Whited (2004) while the assumed depreciation rate reflects an average of 0 for non-decaying stock variables and the 14.5% depreciation rate assumed by Hennessy and Whited for real capital stocks. In this case, attenuation bias exceeds one-half the causal effect if the shock probability (Λ) exceeds 5.26%.

Moving beyond binary settings, suppose b can take on $N \geq 2$ possible values and that all potential transition-to states are equally likely. That is, transitions are uniformly distributed in the sense that $\Pi_{ij} = 1/(N - 1)$ for all $i \neq j$. Suppose also that the shock probability is the same for all N states ($\Lambda_i = \Lambda$ for all i). As shown in the appendix, the next proposition follows from applying equation (12) to this stochastic process.

Proposition 4 *If the shock probability is a constant Λ for all N possible states and shocks are uniformly distributed with $\Pi_{ij} = 1/(N - 1)$ for all $i \neq j$, then the causal effect implied by an observed shock response is given by*

$$CE_{ij}^b = SR_{ij}^b \times \left[1 - \frac{(1 - \delta)\Lambda N/(N - 1)}{r + \delta + (1 - \delta)\Lambda N/(N - 1)} \right]^{-1},$$

with $\text{sgn}(CE_{ij}^b) = \text{sgn}(SR_{ij}^b)$ and $|CE_{ij}^b| > |SR_{ij}^b|$ if $\Lambda > 0$.

The preceding proposition shows that if there is a common shock probability and transitions are uniformly distributed, then there is necessarily bias, and this bias features attenuated shock responses relative to causal effects. Here too, the implied attenuation bias is quite severe. To illustrate, suppose the econometrician inhabits an economy with $r = 2.5\%$ and $\delta = 7.25\%$. And suppose there are 10 possible states. In this case, the absolute value of the shock response will be less than one half that of the causal effect if the shock probability exceeds 9.47%.

As a second example of the preceding proposition, suppose the shock probability is $\Lambda = (N - 1)/N$. In this particular case, the next period state is an i.i.d. random variable with each state, including the present state, occurring with probability $1/N$. Here the absolute value of the causal effect will be an order of magnitude larger than that of the shock response, with $CE = SR(1 + r)/(r + \delta)$. In fact, this is just an example of a more general relationship. To see this, note that if the benefit flow is an i.i.d. random variable, transient q can be expressed as:

$$q_i = \frac{b_i + (1 - \delta)\mathbb{E}[b]/(r + \delta)}{1 + r}. \quad (21)$$

Computing the implied ratio of shock responses to causal effects (equation (17)), and rearranging terms, we have the following proposition.

Proposition 5 *If the benefit flow b is an i.i.d. random variable, then the causal effect implied by an observed shock response is given by*

$$CE_{ij}^b = SR_{ij}^b \times \frac{1 + r}{r + \delta},$$

with $\text{sgn}(CE_{ij}^b) = \text{sgn}(SR_{ij}^b)$ and $|CE_{ij}^b| > |SR_{ij}^b|$.

The preceding proposition shows that with i.i.d. shocks to the benefit flow, there will be attenuation bias. Moreover, this bias will be very severe. Intuitively, with i.i.d. shocks, the current

period’s flow of benefits will constitute a relatively small fraction of the total shadow value. For example, if $r = 2.5\%$ and $\delta = 7.25\%$, then the causal effect will be more than ten times the shock response.

Consider next the case of unanticipated temporary changes to the flow of benefits. In particular, assume state i has the property that the probability of a transition out of that state (Λ_i) is infinitesimal, which implies $q_i \approx q_\infty^i$. However, we place no restriction on this state’s conditional transition-to probabilities, e.g. one can assume $\Pi_{ik} > 0$ for all $k \neq i$. Suppose also that the potential transition-to state j has the property that if a transition from b_i to b_j occurs, the change will be “temporary” in the sense that the benefit variable can only revert back to its initial value b_i without transitioning to any other state. We then have the following proposition.

Proposition 6 *Suppose an observed transition from state i to state j was unanticipated and temporary, with Λ_i infinitesimal and $\Pi_{ji} = 1$. Then the causal effect implied by the observed shock response is*

$$CE_{ij}^b \approx SR_{ij}^b \times \left[1 - \frac{(1 - \delta)\Lambda_j}{r + \delta + (1 - \delta)\Lambda_j} \right]^{-1},$$

with $\text{sgn}(CE_{ij}^b) = \text{sgn}(SR_{ij}^b)$ and $|CE_{ij}^b| > |SR_{ij}^b|$.

A few points are worth noting regarding the previous proposition. First, although the proposition labels the transition from state i to state j as temporary, since $\Pi_{ji} = 1$, the proposition actually holds regardless of the expected state j occupation time Λ_j^{-1} . Second, the bias here is identical in form to the bias in the case of binary shocks, as described in Proposition 3. In particular, again the shock response will have the correct sign but with an attenuated absolute value. Finally, this bias can be severe even if the probability of shocks is small. For example, if $r = 2.5\%$ and $\delta = 7.25\%$, then the shock response is less than one-half the causal effect if the shock probability Λ_j is greater than 10.5%.

Having just analyzed the case of unanticipated temporary shocks, we turn finally to a characterization of bias in the case of partially anticipated permanent shocks. We have the following result.

Proposition 7 *Suppose state i is such that any shock to the benefit flow b will be permanent. If the realized transition is from b_i to b_k , then the causal effect implied by the observed shock response*

is given by

$$CE_{ik}^b = SR_{ik}^b \times \left[1 - \left[1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right) \right]^{-1} \frac{\overbrace{\sum_{j \neq i} \Pi_{ij} (b_j - b_i)}^{\text{Expected Change}}}{\underbrace{b_k - b_i}_{\text{Realized Change}}} \right]^{-1}.$$

The preceding proposition has a number of important implications. First, note that if the expected change in the causal variable is zero, then the observed shock response is equal to its corresponding causal effect. However, if the expected change is not zero, the shock response is necessarily a biased estimator of the theory-implied causal effect. Most importantly, from Proposition 7 it follows that with permanent shocks, there can be attenuation bias, overshooting, and sign reversals. In particular, we have the following corollary delineating bias regions.

Corollary. *Suppose state i is such that any shock to the benefit flow b will be permanent, with the conditional expectation of the change not equal to zero. Then the sign and magnitude of the wedge between causal effects and the shock response depends on the realized value b_k . If the expected change was positive, then*

$$\begin{aligned} b_k - b_i < 0 &\Rightarrow SR_{ik} < CE_{ik} < 0 \\ b_k - b_i \in \left(0, \frac{\sum_{j \neq i} \Pi_{ij} (b_j - b_i)}{1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right)} \right) &\Rightarrow SR_{ik} < 0 < CE_{ik} \\ b_k - b_i \geq \frac{\sum_{j \neq i} \Pi_{ij} (b_j - b_i)}{1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right)} &\Rightarrow 0 \leq SR_{ik} < CE_{ik}. \end{aligned}$$

If the expected change was negative, then

$$\begin{aligned}
b_k - b_i &> 0 \Rightarrow SR_{ik} > CE_{ik} > 0 \\
b_k - b_i &\in \left(\frac{\sum_{j \neq i} \Pi_{ij}(b_j - b_i)}{1 + \Lambda_i^{-1} \left(\frac{r+\delta}{1-\delta} \right)}, 0 \right) \Rightarrow SR_{ik} > 0 > CE_{ik} \\
b_k - b_i &\leq \frac{\sum_{j \neq i} \Pi_{ij}(b_j - b_i)}{1 + \Lambda_i^{-1} \left(\frac{r+\delta}{1-\delta} \right)} \Rightarrow 0 \geq SR_{ik} > CE_{ik}.
\end{aligned}$$

The intuition for the corollary is as follows. Consider, say, a setting in which the conditional expectation of the change in b was positive. If the realized change is negative, the shock response will have the correct (negative) sign but will overshoot the causal effect in absolute value terms. Heuristically, an objectively bad shock becomes a terrible shock if the expected shock was positive. If instead the realized shock involved a large increase in b , the shock response will have the correct positive sign but be biased downwards. Heuristically, objectively great news becomes good news if positive news was expected. Finally, if the realized shock entails a sufficiently small (relative to expectations) increase in b , the shock response will actually have the wrong sign. Heuristically, an objectively good shock becomes a bad shock if a better shock had been expected.

The corollary allows one to readily compute the ex ante probability of the various biases, attenuation, overshooting, and sign reversals, depending on the probability of a shock, the probability distribution of shocks, the discount rate, and the rate of stock decay. One simply needs to compute the probability of the shock falling into the respective bias regions described in the corollary.

To illustrate, suppose $r = 2.5\%$ and $\delta = 7.25\%$. Suppose also that a shock is viewed as relatively unlikely, with $\Lambda_i = 20\%$, with any change in b being uniformly distributed on $[-0.15, 0.85]$. It follows from the corollary that the shock response will overshoot the causal effect if the realized change is negative, a change that occurs with probability 15%. For positive shocks in the interval $[0, 0.23]$, the shock response will actually be negative. That is, there is a 23% probability of the shock response having the wrong sign. Finally, for the remaining 62% of potential shock realizations, the shock response will undershoot the causal effect.

Continuing with the preceding example, Figure 1 plots the ratio of shock response to causal effect under alternative depreciation rates of 0%, 7.25%, and 14.5%. In all cases, the shock response

differs from the causal effect, with overshooting for negative shocks, undershooting for sufficiently large positive shocks, and incorrect signs for small positive shocks. Consistent with the corollary, the measure of the sign reversal region decreases with the depreciation rate, implying the problem of sign reversals is less severe for shorter-lived assets.

III. Identifying Assumptions

In the preceding section it was shown that empirically observed responses to exogenous shocks to the intertemporal benefit flow variable b will not generally recover the theory-implied causal effect. In fact, it was shown that shock responses do not even necessarily have the same sign as the causal effect. However, a general correction algorithm was provided (Proposition 2), along with specific correction formulae for binary settings (Proposition 3), transient uniformly distributed shocks (Proposition 4), i.i.d. shocks (Proposition 5), unanticipated temporary shocks (Proposition 6), and anticipated permanent shocks (Proposition 7). The objective of this section is to provide empiricists with a set of identifying assumptions that would need to be demonstrated beyond exogeneity, ensuring shock responses are equal in magnitude or sign to theory-implied causal effects.

One expects shock responses to approximate causal effects if the shock process features almost completely unanticipated changes that are nearly-permanent. Indeed, we have the following proposition.

Proposition 8 *In the limit as all shock probabilities tend to zero, each measured shock response converges to its respective causal effect.*

The preceding result provides a formal justification for the invocation of “unexpected and permanent shocks” as an identifying assumption for the recovery of theory-implied causal effects. However, it suffers from some practical limitations. First, natural experiments exploiting rare events must be truly short in supply. A second problem associated with claiming to exploit a permanent shock is that one tends to be interested in analyzing a given historical shock response with an eye to making forecasts of future shock responses or with the goal of setting policy optimally in the future, which contradicts the notion that the observed shock is permanent. Conveniently, we have the following alternative identifying assumption.

Proposition 9 *The necessary and sufficient condition for shock responses to equal causal effects for each possible jump in the intertemporal benefit flow is that b is a martingale: the best state (1) and worst state (N) must be absorbing, with remaining states being absorbing or featuring mean-zero changes.*

To provide intuition for the preceding proposition it is convenient to express the value of a unit of stock as the discounted value of the expected flow of benefits it provides. Under the martingale assumption, this value can be written as:

$$\begin{aligned} q_i &= \sum_{t=1}^{\infty} \left(\frac{1}{1+r} \right)^t (1-\delta)^{t-1} \{ \mathbb{E}_0[b(t)|b(0) = b_i] \} \\ &= \sum_{t=1}^{\infty} \left(\frac{1}{1+r} \right)^t (1-\delta)^{t-1} b_i \\ &= \frac{b_i}{r+\delta}. \end{aligned} \tag{22}$$

The second line in the preceding equation holds under the assumption that the benefit flow is a martingale. Effectively, with the martingale assumption, the shadow value of the stock capitalizes the current state as if it will last forever. And this holds before each shock and after each shock. Thus, at each point in time the firm acts *as if* the current state will last forever despite knowing it will not. In such environments, the econometrician can directly recover causal effects from shock responses—and this is true even if shocks occur frequently. Effectively, under the martingale assumption, transient q is equivalent to permanent q_{∞} period-by-period, so that accumulation would be identical across the stochastic and non-stochastic economies.

In some cases, an empiricist may be interested in the more limited objective of avoiding the possibility of a shock response having a different sign than the theory-implied causal effect. A description of identifying assumptions ensuring against sign reversals is therefore worthwhile. Recall, we have derived a number of such sign-identification conditions above. We offer now an alternative sufficient condition for avoiding sign reversals.

Since $b_j > b_i$ for all $j < i$, and accumulation is increasing in q , a sufficient condition for the absence of sign reversals is $q_j \geq q_i$ for all $j < i$. This will necessarily be true under *stochastic monotonicity*: For all $j \neq i$ and $\hat{b} \in \{b_1, \dots, b_N\}$,

$$b_j > b_i \Rightarrow \Pr \left[b_{t+1} \leq \hat{b} | b_t = b_j \right] \leq \Pr \left[b_{t+1} \leq \hat{b} | b_t = b_i \right].$$

That is, under stochastic monotonicity, if $b_j > b_i$, then the distribution of future states conditional upon current state b_j is first order stochastic dominant to the distribution conditional upon current state b_i . In practice, applied researchers might advance a verbal plausibility argument in defense of a stochastic monotonicity assumption. More formally, Lee, Linton and Whang (2009) offer a test statistic for assessing the null hypothesis of stochastic monotonicity given an observed time-series. Alternatively, given an estimated Markov transition matrix, Keilson and Kester (1977) present necessary and sufficient conditions for stochastic monotonicity. Finally, as shown by Keilson and Kester, a number of common processes are stochastically monotone, such as random walks with absorbing or impenetrable boundaries.

The following proposition summarizes the conditions we have derived for the avoidance of sign reversals, but insufficient for correct identification of causal effect magnitudes.

Proposition 10 *For changes in the intertemporal benefit flow b , the shock response has the same sign as the causal effect if: the benefit flow is binary (Proposition 3); shocks are transient and uniformly distributed (Proposition 4); shocks are i.i.d. (Proposition 5); the shock is unanticipated and temporary (Proposition 6); or the benefit process is stochastically monotone. If shocks are permanent with non-zero expected changes, sufficient conditions for the avoidance of sign reversals are that the realized change differs in sign from the expected change or that the realized change is equal in sign but not less than the expected change (Proposition 7).*

IV. External Validity and Extrapolation

Up to this point attention has been confined to determining the relationship between measured shock responses and causal effects. This section considers the related question of how observed shock responses can be extrapolated. In particular, we first consider how an observed shock response in a specific economy can be used to forecast future shock responses in that same economy. Next, we consider how an observed shock response in a given economy can be used to forecast future shock responses in an otherwise identical economy endowed with a different shock generating process.

A. Challenges to External Validity

An observation following directly from the model is that it is not generally valid to use the shock response in one economy to forecast the shock response in another economy. This argument holds

even if the change in the policy variable is the same and the economies have identical production technologies. To see this, note that it follows from equation (12) that if economies A and B have different shock processes, then their augmented transition matrices will differ ($\mathbf{T}^A \neq \mathbf{T}^B$), implying differences in q , and different responses to shocks to the benefit flow b . In other words, as argued by Lucas (1976), if the data generating process is changed (here represented by \mathbf{T}), decision rules (here the SR_{ij}^b) will change. This is a strong limitation on external validity.

In fact, the Lucas Critique is just one limit on extrapolation. The model also illustrates another important limit to external validity. Specifically, even if one considers a single economy and holds fixed the augmented transition matrix \mathbf{T} , so that the Lucas Critique does not apply, it is generally not valid to directly use a previous shock response to forecast future shock responses.

To illustrate these two limits to external validity, we provide two numerical examples featuring shocks to the benefit variable b . Specifically, it is assumed that tax rate, call it τ , can take on ten possible values ranging from a minimum of 0% to a maximum of 45% with the assumed benefit flow taking the form $b = 1 - \tau$. Thus, the tax rate jump across neighboring states is 5 percentage points. Both examples assume that the tax rate can only jump to one of its direct neighbors. Arguably, small yet discrete jumps of the sort considered here are more realistic than infinitesimal tax rate jumps. Finally, both examples use the previous parameter value assumptions, with $r = 2.5\%$ and $\delta = 7.25\%$.

The first example considers an economy in which corporate tax rates have a tendency to drift downward. Specifically, it is assumed that the probability of a decrease in the tax rate is 15% and the probability of an increase is 5%. The shock responses are presented in Table I. As shown, the fact that there is a higher probability of a tax rate reduction results in asymmetric shock responses. In particular, since a tax decrease is more likely, the shadow value of the stock variable (q_i) is pushed upward, implying relatively large (small) responses to tax rate increases (decreases). For example, starting at $\tau = 5\%$, the shock response for a 5 percentage point tax rate decrease is .236 while that for a 5 percentage point tax rate increase is -.364, a 54% difference. In addition, the magnitude of shock responses varies greatly across the different states. For example, the minimum shock response is .236 and the maximum shock response is .503, a difference of 113%.

The previous example highlights the potential pitfalls associated with extrapolating shock responses within a given economy. Consider now the challenge of extrapolating shock responses

across economies. To this end, suppose Table I captures shock responses in Economy 1, with Table II capturing shock responses in an otherwise identical economy, but one where the tax rate has a mean-reverting tendency. Specifically, Table II assumes that for tax rates in the low range, from 0% to 20%, the probability of a tax rate increase is 15% and the probability of a decrease is 5%. Conversely, for tax rates in the high range, from 25% to 45%, the probability of a tax rate increase is 5% and the probability of a tax rate decrease is 15%. The first obvious point to note is that there are large state-by-state differences in shock responses across the two economies. Moreover, there is a systematic difference in the behavior of relative shock responses. In Economy 1 the smallest shock responses are found at the low tax rates. In contrast, in Economy 2, the smallest shock response is found for mid-range tax rates.

Despite the fact that the two economies are identical in terms of the technologies and objective functions of their respective firms, the shock responses in Economy 1 provide little in the way of directly useful information in forecasting shock responses in Economy 2. What is needed for valid extrapolation is a procedure for making adjustments to shock responses to account for differences in data generating processes. The problem of extrapolation is analyzed in the next subsection.

B. Extrapolation and Parameter Inference

This subsection considers the problem of extrapolating shock responses both within and across economies. Consider first the simplest case in which one has observed the response of firms in Economy 1 to a transition in the price of the stock from p_i to p_j . The researcher wants to use the observed shock response to forecast the behavior of firms in this same economy, or to forecast the behavior of firms in another economy facing a different shock generating process. In either case, one can think of the objective as using the shock response SR_{ij}^p to forecast shock responses to a transition in the price of the stock from, say, p_h to p_k . From equation (16) we have:

$$SR_{hk}^p = (-\xi)(p_k - p_h) = \underbrace{\left(\frac{SR_{ij}^p}{p_j - p_i} \right)}_{-\xi} (p_k - p_h) = SR_{ij}^p \times \left[\frac{p_k - p_h}{p_j - p_i} \right]. \quad (23)$$

Notice, the final inequality in the preceding equation tells us that regardless of the underlying stochastic processes, the correct procedure for extrapolating the measured price shock response is to simply scale by the relative size of the shocks. The second inequality in the preceding equation tells us that extrapolation of shock responses implicitly entails inferring the structural parameter ξ

measuring ease of stock adjustment. In fact, it is the implicit/explicit inference of a policy-invariant parameter that allows for extrapolation.

Consider next the problem of extrapolating an observed shock to the cost of stock adjustment ξ either within a given economy or to another economy with a different shock generating process. The objective is to use an observed shock response SR_{ij}^ξ to forecast shock responses to a transition from ξ_h to ξ_k . From equation (16) we have:

$$SR_{hk}^\xi = (q - p)(\xi_k - \xi_h) = \underbrace{\left(\frac{SR_{ij}^\xi}{\xi_j - \xi_i} \right)}_{q-p} (\xi_k - \xi_h) = SR_{ij}^\xi \times \left[\frac{\xi_k - \xi_h}{\xi_j - \xi_i} \right]. \quad (24)$$

Again, the final inequality in the preceding equation tells us that regardless of the underlying stochastic processes, the correct procedure for extrapolating the measured adjustment cost shock response is to simply scale by the relative size of the shocks. And once again, the second inequality informs us that extrapolation can be understood as based upon implicit/explicit structural parameter inference, here the gross gain to accumulation, $q - p$.

Consider finally the problem of extrapolating an observed shock to the intertemporal benefit flow b either within a given economy or to another economy with a different shock generating process. The objective is to use an observed shock response SR_{ij}^b to forecast shock responses to a transition from b_h to b_k . From equation (16) we have:

$$SR_{hk}^b = \xi(q_k - q_h) = \underbrace{\left(\frac{SR_{ij}^b}{q_j - q_i} \right)}_{\xi} (q_k - q_h) = SR_{ij}^b \times \left[\frac{q_k - q_h}{q_j - q_i} \right]. \quad (25)$$

Note there is a superficial resemblance between the extrapolation formula given in equation (25) and those given in equations (23) and (24). However, the latter two formulae entail scaling the observed shock response by the relative size of the shocks to the underlying causal variable, with no need to consider the underlying stochastic processes. In contrast, equation (25) indicates that extrapolating shocks to the intertemporal benefit flow b demands scaling by relative changes in the shadow value of the stock variable. This requires a careful accounting for underlying stochastic processes via equations (12) and (13).

Intuitively, the preceding extrapolation formula can be understood as a two-step process. In the first step, one computes the shock response per unit of shadow value (ξ) by normalizing the

historical shock response by the change in shadow value at the shock date. This is then re-scaled by the change in shadow value associated with the contemplated future-date transition.

V. Numerical Example: Taxes and Corporate Investment

This section provides a numerical example illustrating how one can go about inferring causal effects and causal parameters based upon observed shock responses assuming the observed stochastic process for the causal variable b fails to satisfy one of the previously derived conditions for equality of shock responses and causal effects. The primary objective here is to illustrate how one can apply Proposition 2 in a real-world setting where the data generating process must be inferred based upon an observed time-series. In addition, the numerical example illustrates the potentially large quantitative impact of dynamic tax policy uncertainty.

Consider an econometrician interested in estimating the sign and magnitude of the causal effect of taxes on corporate investment. For the sake of the numerical illustration, we assume the periodic benefit flow per unit of capital is $b_t \equiv 1 - \tau_t$ where τ_t is the observed history of effective tax rates on corporate investment over the period from 1953-2005, as computed by Gravelle (1994) and the Congressional Research Service (2006).⁷

To focus on the challenges arising from dynamic uncertainty, we abstract from concerns about endogeneity here and assume the observed tax rate series was generated by the following exogenous process

$$\tau_t = \frac{1}{1 + \exp(-\varepsilon_t)} \quad (26)$$

where ε_t is an AR(1) process with unconditional mean μ and autocorrelation ρ :

$$\varepsilon_t = \mu(1 - \rho) + \rho\varepsilon_{t-1} + u_t. \quad (27)$$

The shock u_t is assumed to be normally distributed with mean 0 and variance σ^2 . Notice, the posited logistic functional form for the effective tax rate, which has also been used to model factor shares, has the property that the tax rate has support on the proper interval $(0, 1)$.

We map the observed sequence of effective tax rates back to an implied sequence for the AR(1) process by inverting equation (26):

$$\varepsilon_t = -\ln(\tau_t^{-1} - 1).$$

⁷This is a simplification because we do not break the total effective tax rate into its constituent parts.

We then estimate the parameters of the underlying AR(1) process using Maximum Likelihood. We estimate the unconditional mean μ is 0.0158, the autocorrelation ρ is 0.917, and the innovation variance σ^2 is 0.0261.

We next obtain a discrete state-space approximation for the estimated AR(1) process using the method proposed by Adda and Cooper (2003), which is a variation of Tauchen (1986) and Tauchen and Hussey (1991). Discretizing the state space into three points we obtain a tax rate of 35.5% in the best state (state 1), 50.4% in state 2, and 62.7% in state 3. If we instead discretize the state space into four points, we obtain state-contingent tax rates of 33.3%, 45.9%, 54.9%, and 67.4%.

Table III provides details on the estimated transition rates for three and four states. In the case of the three state approximation, the probability of a shock falls between 15% and 25%. In the case of the four state approximation, the probability of a shock falls between 17% and 33%.

We next use equations (12) and (13) to estimate the shadow value of capital (q) in each tax rate state. We then compute the implied ratio of shock responses to causal effects using equation (17) under the assumed parameter values $r = 2.5\%$ and $\delta = 7.25\%$.

Table IV shows the implied ratios of shock responses to theory-implied causal effects for each possible tax rate transition. From equation (17) it follows that two matrices are symmetric. In the case of the three state approximation, the implied ratio of shock responses to causal effects falls in a narrow range from .425 to .427. In the case of the four state approximation, the implied ratio of shock responses to causal effects ranges from .440 to .548. In general, implied shock responses amount to less than one-half the theory implied causal effect.

Clearly, a halving of the true causal effect can lead to erroneous conclusions. For example, one might be lead to conclude that taxes have only a second-order effect on real investment. More importantly, this type of bias could lead to an erroneous halving of the estimated welfare loss from taxation. For example, Rosen (1992) calculates the excess burden from taxation as the difference between taxes collected and the loss to the firm. Applying his method, the per-period excess burden

from corporate taxes on the return to capital would be calculated as:⁸

$$\begin{aligned}
EB &= \frac{1}{2} \times \Delta q \times \Delta a \\
&= \frac{1}{2} \times \tau q_{pretax} \times [\xi(q_{pretax} - p) - \xi((1 - \tau)q_{pretax} - p)] \\
&= \frac{1}{2} \tau^2 q_{pretax}^2 \xi.
\end{aligned} \tag{28}$$

Notice, the excess burden is linear in ξ which is the structural parameter capturing the “elasticity” of investment to changes in the shadow value of capital. If an econometrician were to attempt a calculation of the excess burden, while failing to account for transience in the structural parameter estimate, they would understate the excess burden from taxation by roughly 50%. To see this, note that equation (20) implies that the estimated excess burden here would be:

$$\widehat{EB} = \frac{1}{2} \tau^2 q_{pretax}^2 \widehat{\xi} = \frac{1}{2} \tau^2 q_{pretax}^2 \left(\frac{SR_{ij}^b}{CE_{ij}^b} \right) \xi = EB \times \left(\frac{SR_{ij}^b}{CE_{ij}^b} \right). \tag{29}$$

VI. Alternative Technology Assumptions

Thus far we have focused on the classical linear-quadratic technology in which the firm receives a benefit inflow that is linear in the stock, facing adjustment costs that are quadratic in accumulation. This technology merits focus since, as shown above (equation (7)), it maps directly to the linear causal effects econometric model. Specifically, with such a linear-quadratic technology, theory-implied causal effects are proportional to changes in the underlying causal parameters: benefit inflow (b); price of the stock (p); and ease of stock adjustment (ξ). This section analyzes the relationship between shock responses and theory-implied causal effects for firms endowed with different benefit inflow and adjustment cost functions.

As argued by Summers (1981), in order to empirically replicate observed investment dynamics, it appears necessary to impose large stock adjustment costs. Otherwise predicted investment appears to be excessively sensitive to the shadow value of capital. Nevertheless, for light capital goods, large adjustment costs may be considered unreasonable. With this in mind, we consider first the extreme case of a firm enjoying the ability to adjust the stock variable s without incurring any adjustment costs. Suppose also that the benefit inflow is concave, taking the form $bF(s)$, where b

⁸See Rosen (1992) page 327.

is a stochastic process and where F is a strictly increasing concave function satisfying the Inada conditions.⁹ As shown in the appendix, the optimal policy at each point in time t is to increase the stock until its discounted marginal product, inclusive of expected resale value, is just equal to its current market price:¹⁰

$$\begin{aligned} p_t &= \left(\frac{1}{1+r} \right) [b_t F'(s_t^*) + (1-\delta)\mathbb{E}_t[p_{t+1}]] \\ \Rightarrow s_t^* &= [F']^{-1} \left[\frac{(1+r)p_t - (1-\delta)\mathbb{E}_t[p_{t+1}]}{b_t} \right]. \end{aligned} \quad (30)$$

By way of comparison, in the context of a deterministic dynamic model in which b and p are treated as parameters rather than processes, the optimal stock level can be expressed in terms of the Jorgenson (1963) user cost of capital, $(r+\delta)p$. In particular:

$$s^{**} = [F']^{-1} \left[\frac{(1+r)p - (1-\delta)p}{b} \right] = [F']^{-1} \left[\frac{(r+\delta)p}{b} \right]. \quad (31)$$

From equation (30) it follows that here the firm's response to price shocks will depend upon the law of motion for the price. Intuitively, with frictionless stock adjustment, expected price dynamics matter since the firm acts as if it will sell its asset stock at the end of each period and then re-optimize. Comparing equation (30) with equation (31) it is also apparent that with frictionless stock adjustment, responses to price shocks do not generally equal theory-predicted causal effects. However, note that if the stock price is a martingale, or if price shocks are completely unanticipated and permanent, then $s^* = s^{**}$ and price shock responses directly recover causal effects.

Equation (30) also reveals that with frictionless stock adjustment, the firm's response to benefit shocks are actually invariant to the law of motion for the benefit inflow scalar b . Here benefit shock responses equal theory-predicted causal effects. Intuitively, with frictionless stock adjustment, the firm re-optimizes the level of the stock each period based on the currently observed inflow of benefits.

Summarizing, we have the following proposition.

Proposition 11 [*Concave Benefit Inflow and No Stock Adjustment Costs*]: *With exogenous changes in the spot price of the stock, shock responses recover theory-implied causal effects only if the price is a martingale or the shock is completely unanticipated and permanent. With exogenous changes*

⁹If profits were linear here, the firm would either buy or sell an infinite amount of the stock.

¹⁰The derivations in this section do not impose Markov shocks.

in the benefit inflow scalar, measured shock responses directly recover theory-implied causal effects regardless of the underlying stochastic process.

Recall, in our baseline technology with linear benefits and quadratic adjustment costs, price shock responses directly recover causal effects, while benefit shock responses only do so if benefits are a martingale or if benefit shocks are completely unanticipated and permanent. In contrast, the preceding proposition shows that with concave benefits and zero adjustment costs, benefit shock responses directly recover causal effects, while price shock responses only do so if the stock price is a martingale or if price shocks are completely unanticipated and permanent.

Consider next a firm facing quadratic adjustment costs ($a^2/2\xi$), as in the baseline model, but suppose as immediately above that the benefit inflow is concave in the stock, taking the form $bF(s)$. Comparative statics do not suffice to evaluate causal effects in this setting since, even with zero uncertainty, the optimal accumulation policy (a) would vary over time depending on the level of the capital stock relative to the steady state. In particular, as in the model of Summers (1981), a completely unanticipated and permanent change in parameters would be modeled as a jump onto a new saddle-point path converging to a new steady state. For example, starting at an initial steady state, an increase in b (decrease in p) would be associated with an increase in accumulation converging to a higher steady state stock. Nevertheless, it is apparent that a correct interpretation of real-world shocks to either the stock price or the benefit flow would here require correct modeling of the stochastic process generating shocks. In particular, as shown in the appendix, with shocks to the stock price and/or the benefit flow, the optimal accumulation policy under the posited technology takes the following form:

$$a_t^* = \xi \left[\left(\frac{1}{1+r} \right) \left(\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^\tau \mathbb{E}_t[b_{t+\tau} F'(s_{t+\tau})] \right) - p_t \right]. \quad (32)$$

The key implication of the preceding equation is that if the benefit inflow function is concave, responses to both price and benefit shocks will depend upon the data generating process for the shocks. To see this, note that optimal accumulation at date t depends upon expectations of the future path of the stock which will depend upon expectations regarding the future path of prices and benefits.

Finally, let us again consider a firm with a concave benefit function but assume now that adjustment costs depend upon the level of the stock. In particular, following Summers (1981), assume

adjustment costs take the form $\gamma a_t^2/2s_{t-1}$. As shown in the appendix, under such an adjustment cost function, the optimal rate of accumulation takes the following form:

$$\frac{a_t^*}{s_{t-1}} = \frac{1}{\gamma} \left[\left(\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^\tau \mathbb{E}_t \left\{ b_{t+\tau} F'(s_{t+\tau}) + \frac{1}{1+r} \frac{\gamma}{2} \left(\frac{a_{t+\tau+1}}{s_{t+\tau}} \right)^2 \right\} \right) - p_t \right]. \quad (33)$$

The key implication of the preceding equation is that with stock-dependent adjustment costs, responses to both benefit shocks and price shocks will depend upon the data generating process for the shocks. Moreover, this is true even if the benefit function is linear. To see this, note that the optimal accumulation rate at date t depends upon expectations of the future accumulation rate squared which will depend upon expected future prices. Intuitively, under the posited adjustment cost function, installed capital provides a direct inflow of benefits as well as an indirect benefit in terms of reducing future adjustment costs. In turn, the present value of the reduction in future adjustment costs depends upon expected future prices.

VII. Conclusions

The first objective of this paper is to highlight the existence of a large and important gap between underlying theories and the types of empirical evidence often used to assess the validity or importance of those theories.

On one hand modelers typically derive testable implications of underlying theories using comparative statics analysis. Heuristically, comparative statics yields predictions regarding how agents will respond to completely unanticipated completely permanent shocks. On the other hand, empiricists routinely exploit data generating processes that differ markedly from the type of rare event shocks that comparative statics analysis approximates, e.g. state-level anti-takeover legislation or tax changes. Indeed, if empiricists were to confine attention to truly rare events, they would need to wait a rather long time for nature to provide experiments shedding light on important economic questions. Moreover, the fact that empirical evidence is often used to inform policy decisions or to make predictions about responses to future shocks contradicts the assumption that exploited shocks are permanent.

Deaton (2010) expresses reservations about instrumental variables estimation, and specifically the lack of an ex ante characterization of what such estimation is credibly recovering, with estimates

varying with the choice amongst valid instruments: “This goes beyond the old story of looking for an object where the light is strong enough to see; rather, we have at least some control over the light but choose to let it fall where it may and then proclaim that whatever it illuminates is what we were looking for all along.” We have shown that a similar problem of interpretation emerges with shock-based inference in dynamic environments. In general, shock responses do not directly recover theory-implied causal effects (comparative statics). Moreover shock-based estimates are not generally comparable to each other, with estimates varying within and across alternative data generating processes.

What is to be done? Wolpin (2013) and Rust (2014) argue for the general need to filter empirical evidence through the prism of models, a point to which many pay lip service. The question is how to make this prescription operational in a tractable way. The approach we have taken here straddles reduced-form and structural approaches. On one hand, we focus on the best type of reduced-form evidence—evidence that exploits exogenous shocks and thus avoids selection and endogeneity bias. On the other hand, we build a general stylized model mimicking the data generating process exploited by the econometrician in order to flush out the economic meaning of broad classes of reduced-form evidence derived in dynamic settings.

The second and primary point of our paper is constructive, showing how empiricists can address head-on generally neglected issues associated with dynamic uncertainty, economic interpretability, and external validity. In particular, we show how once the data generating process has been estimated, observed shock responses can be mapped back to theory-implied causal effects, or generalized within or across shock generating processes. In addition, we derive a battery of auxiliary identifying assumptions, beyond exogeneity, that must be established in order to demonstrate lack of bias in magnitudes/signs. Finally, we offer tractable bias adjustment formulae for various classes of data generating processes.

Our analysis also highlights the critical role played by alternative technological assumptions. For example, with linear benefits and quadratic adjustment costs, the stochastic process matters for interpreting benefit shocks but not for shocks to the price of the stock variable. With concave benefits and zero adjustment costs, these conclusions are reversed. Thus, it is apparent that interpreting shock responses requires making joint assumptions about: orthogonality; data generating processes; and underlying real technologies. The first of these assumptions is often discussed,

but the latter two are generally neglected. Our paper makes a first attempt at dealing with the latter two problems, with heavy emphasis on the linear-quadratic economy. Future work should attempt to provide reduced-form empiricists with analogous tractable methods and formulae for bias adjustment under alternative technological assumptions.

Appendix: Proofs and Derivations

A. Appendix Constant Policy Model Solution

The Bellman equation is:

$$V(s) = \max_a -pa - \frac{1}{2}\xi^{-1}a^2 + \left(\frac{1}{1+r}\right) [bs' + V(s')].$$

subject to

$$s' = (1 - \delta) + a.$$

We conjecture the value function is of the form:

$$V(s) = \hat{q}s + G.$$

Notice, \hat{q} represents the value of a unit of capital already installed at the beginning of the period— which will depreciate during the period. The Bellman equation can then be written as:

$$\hat{q}s + G = -pa - \frac{1}{2}\xi^{-1}a^2 + \left(\frac{1}{1+r}\right) \{G + (\hat{q} + b)[a + (1 - \delta)s]\}.$$

Terms scaled by s must equate so we have:

$$\hat{q} = \left(\frac{1 - \delta}{1 + r}\right) (\hat{q} + b) \Rightarrow \hat{q} = \frac{b(1 - \delta)}{r + \delta}.$$

Consider next optimal accumulation. We must solve:

$$\max_a \left(\frac{1}{1+r}\right) (b + \hat{q})a - pa - \frac{1}{2}\xi^{-1}a^2.$$

To capture the value of capital installed during the period, we write

$$q \equiv \frac{\hat{q}}{1 - \delta}.$$

The investment program can then be written as:

$$\max_a qa - pa - \frac{1}{2}\xi^{-1}a^2 \Rightarrow a^{**} = \xi(q - p).$$

Consider finally growth option value. We know that at the optimal policy:

$$qa - pa - \frac{1}{2}\xi^{-1}a^2 = \frac{1}{2}\xi(q - p)^2.$$

Substituting the preceding equation into the Bellman equation and ignoring the terms scaled by s (which cancel) we obtain:

$$G = \frac{1}{2}\xi(q-p)^2 + \left(\frac{1}{1+r}\right)G \Rightarrow G = \left(1 + \frac{1}{r}\right)\frac{1}{2}\xi(q-p)^2.$$

B. Appendix Shock Model Solution

The Bellman equation is:

$$V_i(s) = \max_a -p_i a - \frac{1}{2}\xi_i^{-1}a^2 + \left(\frac{1}{1+r}\right) \left[b_i s' + (1 - \Lambda_i)V_i(s') + \Lambda_i \sum_{j \neq i} \Pi_{ij} V_j(s') \right]$$

subject to

$$s' = (1 - \delta) + a.$$

We conjecture the value function is of the form:

$$V_i(s) = \hat{q}_i s + G_i.$$

Note that \hat{q}_i captures the value of a unit of capital already installed at the beginning of the period— which will depreciate during the period. The Bellman equation can then be written as:

$$\begin{aligned} \hat{q}_i s + G_i &= -p_i a - \frac{1}{2}\xi_i^{-1}a^2 + \left(\frac{1}{1+r}\right) \left[(1 - \Lambda_i)G_i + \Lambda_i \sum_{j \neq i} \Pi_{ij} G_j \right] \\ &\quad + \left(\frac{1}{1+r}\right) \left[b_i + (1 - \Lambda_i)\hat{q}_i + \Lambda_i \sum_{j \neq i} \Pi_{ij} \hat{q}_j \right] [a + (1 - \delta)s]. \end{aligned}$$

Terms scaled by s in the preceding equation must equate so we have:

$$\hat{q}_i = \left(\frac{1 - \delta}{1 + r}\right) \left[b_i + (1 - \Lambda_i)\hat{q}_i + \Lambda_i \sum_{j \neq i} \Pi_{ij} \hat{q}_j \right].$$

And so we have a linear system for the value of the begin-of-period stock:

$$\hat{q}_i [r + \delta + (1 - \delta)\Lambda_i] - (1 - \delta)\Lambda_i \sum_{j \neq i} \Pi_{ij} \hat{q}_j = (1 - \delta)b_i.$$

Using the augmented transition matrix (equation (13)), the preceding linear system can be written as:

$$\begin{bmatrix} \widehat{q}_1 \\ \dots \\ \widehat{q}_N \end{bmatrix} = [\mathbf{T}^{-1}] \begin{bmatrix} (1 - \delta)b_1 \\ \dots \\ (1 - \delta)b_N \end{bmatrix}.$$

Consider next optimal accumulation. From the Bellman equation it follows we must solve:

$$\max_a \left(\frac{1}{1+r} \right) \left[b_i + (1 - \Lambda_i)\widehat{q}_i + \Lambda_i \sum_{j \neq i} \Pi_{ij}\widehat{q}_j \right] a - p_i a - \frac{1}{2}\xi_i^{-1}a^2.$$

To capture the value of newly-installed units of the stock, we let

$$q_i \equiv \frac{\widehat{q}_i}{1 - \delta}.$$

Our program can be written as:

$$\max_a q_i a - p_i a - \frac{1}{2}\xi_i^{-1}a^2 \Rightarrow a_i^* = \xi_i(q_i - p_i).$$

And as stated in the text we have the following solution for the value of a unit of new accumulation:

$$\begin{bmatrix} q_1 \\ \dots \\ q_N \end{bmatrix} = [\mathbf{T}^{-1}] \begin{bmatrix} b_1 \\ \dots \\ b_N \end{bmatrix}.$$

Consider next the growth option value. Under optimal accumulation:

$$q_i a - p_i a - \frac{1}{2}\xi_i^{-1}a^2 = \frac{1}{2}\xi_i(q_i - p_i)^2.$$

Substituting the preceding equation into the Bellman system and ignoring the terms scaled by s (which cancel), we obtain:

$$G_i = \left(\frac{1}{1+r} \right) \left[(1 - \Lambda_i)G_i + \Lambda_i \sum_{j \neq i} \Pi_{ij}G_j \right] + \frac{1}{2}\xi_i(q_i - p_i)^2.$$

Rearranging terms we obtain the following linear system:

$$\begin{bmatrix} G_1 \\ \dots \\ G_N \end{bmatrix} = [\mathbf{T}_G^{-1}] \begin{bmatrix} \frac{1}{2}(1+r)\xi_1(q_1 - p_1)^2 \\ \dots \\ \frac{1}{2}(1+r)\xi_N(q_N - p_N)^2 \end{bmatrix}$$

where

$$\mathbf{T}_G \equiv \begin{bmatrix} r + \Lambda_1 & -\Lambda_1 \Pi_{12} & \dots & -\Lambda_1 \Pi_{1N} \\ -\Lambda_2 \Pi_{21} & r + \Lambda_2 & \dots & -\Lambda_2 \Pi_{2N} \\ \dots & \dots & \dots & \dots \\ -\Lambda_N \Pi_{N1} & -\Lambda_N \Pi_{N2} & \dots & r + \Lambda_N \end{bmatrix}.$$

C. Appendix Proof of Proposition 4

For brevity, let $R \equiv r + \delta$. We have the following $N \times N$ augmented transition matrix:

$$\mathbf{T} \equiv \begin{bmatrix} R + (1 - \delta)\Lambda & -(1 - \delta)\Lambda/(N - 1) & \dots & -(1 - \delta)\Lambda/(N - 1) \\ -(1 - \delta)\Lambda/(N - 1) & R + (1 - \delta)\Lambda & \dots & -(1 - \delta)\Lambda/(N - 1) \\ \dots & \dots & \dots & \dots \\ -(1 - \delta)\Lambda/(N - 1) & -(1 - \delta)\Lambda/(N - 1) & \dots & R + (1 - \delta)\Lambda \end{bmatrix}.$$

The shadow values are:

$$\begin{bmatrix} q_1 \\ q_2 \\ \dots \\ q_N \end{bmatrix} = \mathbf{T}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_N \end{bmatrix}.$$

where

$$\mathbf{T}^{-1} = \frac{1}{R[R + (1 - \delta)\Lambda N/(N - 1)]} \begin{bmatrix} R + (1 - \delta)\Lambda/(N - 1) & (1 - \delta)\Lambda/(N - 1) & \dots & (1 - \delta)\Lambda/(N - 1) \\ (1 - \delta)\Lambda/(N - 1) & R + (1 - \delta)\Lambda/(N - 1) & \dots & (1 - \delta)\Lambda/(N - 1) \\ \dots & \dots & \dots & \dots \\ (1 - \delta)\Lambda/(N - 1) & (1 - \delta)\Lambda/(N - 1) & \dots & R + (1 - \delta)\Lambda/(N - 1) \end{bmatrix}.$$

Taking differences across rows, the transition response of the shadow value is:

$$q_j - q_i = \left[1 - \frac{(1 - \delta)\Lambda N/(N - 1)}{r + \delta + (1 - \delta)\Lambda N/(N - 1)} \right] \left(\frac{b_j - b_i}{r + \delta} \right).$$

Thus the change in transient q is the stated multiple of the change in permanent q . The result then follows from the fact that $\Delta a = \xi \Delta q$. ■

D. Appendix Proof of Proposition 6

Under the assumptions stated in the proposition, the transition-to state shadow value takes the form:

$$\begin{aligned} q_j &= \left(\frac{1}{1+r} \right) [b_j + (1-\delta)(1-\Lambda_j)q_j + (1-\delta)\Lambda_j q_i] \\ \Rightarrow q_j &= \frac{b_j + (1-\delta)\Lambda_j q_i}{r + \delta + (1-\delta)\Lambda_j}. \end{aligned}$$

Using the fact that $q_i \approx q_\infty^i$, we have:

$$\begin{aligned} q_j &\approx \frac{b_j}{r + \delta + (1-\delta)\Lambda_j} + \left[\frac{(1-\delta)\Lambda_j}{r + \delta + (1-\delta)\Lambda_j} \right] q_\infty^i \\ &= \frac{b_j}{r + \delta} - \left[\frac{(1-\delta)\Lambda_j}{r + \delta + (1-\delta)\Lambda_j} \right] (q_\infty^j - q_\infty^i) \\ &= q_\infty^j - \left[\frac{(1-\delta)\Lambda_j}{r + \delta + (1-\delta)\Lambda_j} \right] (q_\infty^j - q_\infty^i). \end{aligned}$$

It then follows that:

$$\begin{aligned} \frac{CE_{ij}^b}{SR_{ij}^b} &= \frac{\xi(\Delta q_\infty)}{\xi(\Delta q)} = \frac{(b_j - b_i)/(r + \delta)}{q_j - q_i} \\ \Rightarrow CE_{ij}^b &= SR_{ij}^b \times \frac{(b_j - b_i)/(r + \delta)}{q_j - q_i} \\ &= SR_{ij}^b \times \left[1 - \left(\frac{(1-\delta)\Lambda_j}{r + \delta + (1-\delta)\Lambda_j} \right) \right]^{-1}. \blacksquare \end{aligned}$$

E. Appendix Proof of Proposition 7

Suppose the state is i and all possible transition-to states are absorbing. From equation (7) it follows:

$$SR_{ik}^b = CE_{ik}^b + \xi \left[\frac{b_i}{r + \delta} - q_i \right].$$

Further, since all possible transition-to states are absorbing, equation (11) can be written as:

$$\begin{aligned} q_i &= \left(\frac{1}{1+r} \right) \left[b_i + (1-\delta) \left((1-\Lambda_i)q_i + \Lambda_i \sum_{j \neq i} \Pi_{ij} \frac{b_j}{r + \delta} \right) \right] \\ \Rightarrow q_i &= \frac{b_i}{r + \delta + \Lambda_i(1-\delta)} + \frac{(1-\delta)\Lambda_i}{(r + \delta)[r + \delta + \Lambda_i(1-\delta)]} \sum_{j \neq i} \Pi_{ij} b_j \\ \Rightarrow q_i &= \frac{b_i}{r + \delta} + \frac{(1-\delta)\Lambda_i}{(r + \delta)[r + \delta + (1-\delta)\Lambda_i]} \sum_{j \neq i} \Pi_{ij} (b_j - b_i) \\ \Rightarrow q_i &= \frac{b_i}{r + \delta} + (r + \delta)^{-1} \left[1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1-\delta} \right) \right]^{-1} \sum_{j \neq i} \Pi_{ij} (b_j - b_i). \end{aligned}$$

From the preceding expression for q_i it follows:

$$\begin{aligned}
SR_{ik}^b &= CE_{ik}^b - \xi(r + \delta)^{-1} \left[1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right) \right]^{-1} \sum_{j \neq i} \Pi_{ij}(b_j - b_i) \\
&= CE_{ik}^b - \xi \left[\frac{b_k - b_i}{r + \delta} \right] \left[1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right) \right]^{-1} \frac{\sum_{j \neq i} \Pi_{ij}(b_j - b_i)}{b_k - b_i} \\
&= CE_{ik}^b \left[1 - \left[1 + \Lambda_i^{-1} \left(\frac{r + \delta}{1 - \delta} \right) \right]^{-1} \frac{\sum_{j \neq i} \Pi_{ij}(b_j - b_i)}{b_k - b_i} \right]. \blacksquare
\end{aligned}$$

F. Appendix Proof of Proposition 8

Since T is Strictly Diagonal Dominant it is invertible. Moreover, each entry in the inverse is the quotient of a polynomial and the determinant. Thus, fixing, say, the Euclidean norm, the inverse function is continuous on the set of invertible matrices. It follows:

$$\lim_{\Lambda \downarrow \mathbf{0}} [\mathbf{T}(\Lambda)]^{-1} = [\mathbf{T}(\mathbf{0})]^{-1} = (r + \delta)^{-1} \mathbf{I}_N.$$

This implies each q_i converges to $b_i/(r + \delta)$ implying each SR_{ij}^b converges to each CE_{ij}^b . ■

G. Appendix Proof of Proposition 9

Consider an array of distinct states 1 to N recalling the indexing convention entails $b_1 > \dots > b_N$. A necessary and sufficient condition for each treatment response to equal its respective causal effect is that the difference between state-contingent shadow values is equal to their respective difference in shadow values under permanent b values. That is, there must then exist some constant κ such that for all i :

$$q_i = \frac{b_i}{r + \delta} + \kappa.$$

Under this functional form, the equilibrium conditions (11) can be stated as:

$$\begin{aligned}
(r + \delta)\kappa &= \left(\frac{1 - \delta}{r + \delta} \right) \Lambda_1 \sum_{j \neq 1} \Pi_{1j}(b_j - b_1) \\
(r + \delta)\kappa &= \left(\frac{1 - \delta}{r + \delta} \right) \Lambda_2 \sum_{j \neq 2} \Pi_{2j}(b_j - b_2) \\
&\dots \\
(r + \delta)\kappa &= \left(\frac{1 - \delta}{r + \delta} \right) \Lambda_{N-1} \sum_{j \neq N-1} \Pi_{N-1j}(b_j - b_{N-1}) \\
(r + \delta)\kappa &= \left(\frac{1 - \delta}{r + \delta} \right) \Lambda_N \sum_{j \neq N} \Pi_{Nj}(b_j - b_N).
\end{aligned}$$

Next note that any solution to the preceding system entails $\kappa = 0$ since the right-side of the first equation is non-positive while the right-side of the last equation is non-negative. It follows that any candidate solution to the system entails $\Lambda_1 = \Lambda_N = 0$. Further, it must be the case that for intermediate states, either $\Lambda_i = 0$ or the expected change in the benefit flow is zero. ■

H. Appendix Proof of Proposition 11

The program is:

$$\max_{\{s_t\}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \left(\frac{1}{1+r} \right)^{\tau} \{b_{t+\tau-1}F(s_{t+\tau-1}) - [p_{t+\tau}s_{t+\tau} - (1-\delta)s_{t+\tau-1}]\}.$$

The first-order condition with respect to s_t is:

$$\left(\frac{1}{1+r} \right) b_t F'(s_t) = U_t \equiv p_t - \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t[p_{t+1}].$$

Thus, the optimal stock is:

$$s_t^* = [F']^{-1}[(1+r)U_t/b_t]. \blacksquare$$

I. Appendix Solution Quadratic Adjustment Costs and Concave Profits Model

The program is:

$$\max_{\{s_t\}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \left(\frac{1}{1+r} \right)^{\tau} \left\{ \begin{array}{l} b_{t+\tau-1}F(s_{t+\tau-1}) - p_{t+\tau}[s_{t+\tau} - (1-\delta)s_{t+\tau-1}] \\ - \frac{\xi^{-1}}{2} [s_{t+\tau} - (1-\delta)s_{t+\tau-1}]^2 \end{array} \right\}.$$

The first-order condition with respect to s_t is:

$$\left(\frac{1}{1+r} \right) b_t F'(s_t) - p_t - \xi^{-1}a_t + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [p_{t+1} + \xi^{-1}a_{t+1}] = 0.$$

The optimal accumulation policy is:

$$a_t = \xi \left[\frac{b_t F'(s_t)}{1+r} - U_t \right] + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [a_{t+1}].$$

Iterating and applying the law of iterated expectations we have

$$\mathbb{E}_t[a_{t+1}] = \xi \mathbb{E}_t \left[\left(\frac{b_{t+1}F'(s_{t+1})}{1+r} - U_{t+1} \right) \right] + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [a_{t+2}].$$

Substituting back into the accumulation equation we obtain

$$a_t = \xi \left[\frac{b_t F'(s_t)}{1+r} - U_t + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t \left(\frac{b_{t+1}F'(s_{t+1})}{1+r} - U_{t+1} \right) \right] + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [a_{t+2}].$$

Continuing to iterate in this fashion we obtain:

$$a_t = \xi \sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^{\tau} \mathbb{E}_t \left[\frac{b_{t+\tau} F'(s_{t+\tau})}{1+r} - U_{t+\tau} \right].$$

Next note that

$$\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^{\tau} \mathbb{E}_t [U_{t+\tau}] = p_t.$$

It follows that

$$a_t = \xi \left[\left(\frac{1}{1+r} \right) \left(\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^{\tau} \mathbb{E}_t [b_{t+\tau} F'(s_{t+\tau})] \right) - p_t \right].$$

J. Solution Stock-Dependent Adjustment Costs and Concave Profits Model

The program is:

$$\max_{\{s_t\}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \left(\frac{1}{1+r} \right)^{\tau} \left[\begin{array}{l} b_{t+\tau} F(s_{t+\tau}) - p_{t+\tau} [s_{t+\tau} - (1-\delta)s_{t+\tau-1}] \\ - \frac{\gamma}{2} [s_{t+\tau} - (1-\delta)s_{t+\tau-1}]^2 s_{t+\tau-1}^{-1} \end{array} \right].$$

Letting $\alpha \equiv a/s$, the first-order condition for s_t is:

$$p_t + \gamma \alpha_t = b_t F'(s_t) + \left(\frac{1}{1+r} \right) \mathbb{E}_t \left[(1-\delta)p_{t+1} + (1-\delta)\gamma \alpha_{t+1} + \frac{\gamma}{2} \alpha_{t+1}^2 \right].$$

Rearranging terms, the optimal accumulation rate can be expressed in terms of the user cost of capital U :

$$\alpha_t = \frac{1}{\gamma} [b_t F'(s_t) - U_t] + \frac{1}{2(1-\delta)} \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [\alpha_{t+1}^2] + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [\alpha_{t+1}].$$

Iterating on the accumulation equation and applying the law of iterated expectations we have:

$$\begin{aligned} \mathbb{E}_t [\alpha_{t+1}] &= \frac{1}{\gamma} \mathbb{E}_t [b_{t+1} F'(s_{t+1}) - U_{t+1}] \\ &\quad + \frac{1}{2(1-\delta)} \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [\alpha_{t+2}^2] + \left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [\alpha_{t+2}]. \end{aligned}$$

Substituting this term into the optimality condition we obtain:

$$\begin{aligned} \alpha_t &= \frac{1}{\gamma} \mathbb{E}_t \left[\begin{array}{l} [b_t F'(s_t) - U_t] + \left(\frac{1-\delta}{1+r} \right) [b_{t+1} F'(s_{t+1}) - U_{t+1}] \\ \frac{\gamma}{2(1-\delta)} \left(\left(\frac{1-\delta}{1+r} \right) \mathbb{E}_t [\alpha_{t+1}^2] + \left(\frac{1-\delta}{1+r} \right)^2 \mathbb{E}_t [\alpha_{t+2}^2] \right) \end{array} \right] \\ &\quad + \left(\frac{1-\delta}{1+r} \right)^2 \mathbb{E}_t [\alpha_{t+2}]. \end{aligned}$$

Iterating in this fashion, we obtain:

$$\alpha_t = \frac{1}{\gamma} \left[\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^\tau \mathbb{E}_t \left(b_{t+\tau} F'(s_{t+\tau}) - U_{t+\tau} + \frac{1}{1+r} \frac{\gamma}{2} \alpha_{t+\tau+1}^2 \right) \right].$$

Canceling the offsetting terms involving the user cost of capital, the optimal accumulation policy is

$$\alpha_t = \frac{1}{\gamma} \left[\left(\sum_{\tau=0}^{\infty} \left(\frac{1-\delta}{1+r} \right)^\tau \mathbb{E}_t \left\{ b_{t+\tau} F'(s_{t+\tau}) + \frac{1}{1+r} \frac{\gamma}{2} \alpha_{t+\tau+1}^2 \right\} \right) - p_t \right].$$

References

- [1] Aaron, Henry J., 1990, Lessons for Tax Reform, in J. Slemrod Ed., *Do Taxes Matter? The Impact of the Tax Reform Act of 1986*, MIT Press.
- [2] Abel, Andrew, 1982, Dynamic Effects of Permanent and Temporary Tax Policies in a Q Model of Investment, *Journal of Monetary Economics* (9), 353-373.
- [3] Abel, Andrew and Janice Eberly, 1994, A Unified Model of Investment under Uncertainty, *American Economic Review*, 100-128.
- [4] Abel, Andrew and Janice Eberly, 1997, An Exact Solution for the Investment and Value of a Firm Facing Uncertainty, Adjustment Costs, and Irreversibility, *Journal of Economic Dynamics and Control* (21), 831-852.
- [5] Adda, Jerome and Russell W. Cooper, 2003, *Dynamic Economics: Quantitative Methods and Applications*, MIT Press.
- [6] Alesina, Alberto, Silvia Ardagna, Giuseppe Nicoletti, and Fabio Schiantarelli, 2005, Regulation and Investment, *Journal of the European Economic Association*, 791-825.
- [7] Angrist, Joshua D. and Jorn-Steffen Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- [8] Angrist, Joshua D. and Jorn-Steffen Pischke, 2010, The Credibility Revolution in Economics: How Better Research Design is Taking the Con out of Econometrics, *Journal of Economic Perspectives* (24), 3-30.
- [9] Atanassov, Julian, 2013, Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting, *Journal of Finance* (68), 1097-1131.
- [10] Auerbach, Alan J., 1986, Tax Reform and Adjustment Costs: The Impact on Investment and Market Value, *International Economic Review* (30), 831-852.
- [11] Auerbach, Alan J. and James R. Hines, Jr., 1988, Investment Tax Incentives and Frequent Tax Reforms, *American Economic Review* (78), 211-216.

- [12] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, 2015, The Miracle of Microfinance? Evidence from a Randomized Evaluation, *American Economic Journal: Applied Economics*, 22-53.
- [13] Bertrand, Marianne, and Sendhil Mullainathan, 1999, Is There Discretion in Wage Setting? A Test Using Takeover Legislation, *RAND Journal of Economics*, 535-554.
- [14] Bertrand, Marianne, and Sendhil Mullainathan, 2003, Enjoying the Quiet Life: Corporate Governance and Managerial Preferences, *Journal of Political Economy* (111), 1043-1073.
- [15] Bowen, Donald E., Laurent Fresard, and Jerome Taillard, 2016, What's Your Identification Strategy? Innovation in Corporate Finance Research, *Management Science*, 2529-2548.
- [16] Chetty, Raj, 2012, Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply, *Econometrica* (80), 969-1018.
- [17] Congressional Research Service, 2006, Historical Effective Tax Rates on Capital Income, CRS Report 1408.
- [18] Crepon, Bruce, Florencia Devoto, Esther Duflo, and William Pariente, 2013, Estimating the Impact of Microcredit on Those Who Take it Up: Evidence from a Randomized Experiment, 2015, *American Economic Journal: Applied Economics*, 123-150.
- [19] Cummins, Jason G., Kevin A. Hassett, and R. Glenn Hubbard, 1994, A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments, *Brookings Papers on Economic Activity* (2), 1-74.
- [20] Deaton, Angus, 2010, Instruments, Randomization, and Learning about Development, 2010, *Journal of Economic Literature* (48), 424-455.
- [21] Deschenes, Olivier and Michael Greenstone, 2007, Economic Impacts of Climate Change: Evidence from Random Weather Fluctuations, *American Economic Review*.
- [22] Gan, Jie, 2007, Collateral, Debt Capacity and Corporate Investment: Evidence from a Natural Experiment, *Journal of Financial Economics* (85), 709-754.

- [23] Garvey, Gerald and Gordon Hanka, 1999, Capital Structure and Corporate Control: The Effect of Antitakeover Statutes on Firm Leverage, *Journal of Finance* (54), 519-546.
- [24] Givoly, Dan, Carla Hayn, Aharon Ofer, and Oded Sarig, 1992, Taxes and Capital Structure: Evidence from Firms' Response to the Tax Reform Act of 1986, *Review of Financial Studies* (5), 331-355.
- [25] Gourio, Francois and Jianjun Miao, 2008, Transitional Dynamics of Dividend and Capital Gains Tax Cuts, *Review of Economic Dynamics* (14), 368-383.
- [26] Gravelle, Jane, 1994, *The Economic Effects of Taxing Capital Income*, MIT Press.
- [27] Greenstone, Michael, 2002, The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and Census of Manufacturers, *Journal of Political Economy* (110).
- [28] Hassett, Kevin A. and Gilbert Metcalf, 1999, Investment with Uncertain Tax Policy: Does Random Taxation Discourage Investment, *Economic Journal* (109), 372-393.
- [29] Heckman, James J., 1997, Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations, *Journal of Human Resources* (32), 441-462.
- [30] Heckman, James J., 2000, Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective, *Quarterly Journal of Economics* 115 (1), 45-97.
- [31] House, Christopher L., and Matthew D. Shapiro, 2008, Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation, *American Economic Review* (98), 737-768.
- [32] Jorgenson, Dale, 1963, Capital Theory and Investment Behavior, *American Economic Review* (53), 247-259.
- [33] Karpoff, Jonathan, and Michael Wittry, 2015, Institutional and Political Economy Considerations in Natural Experiments: The Case of State Antitakeover Laws, working paper, SSRN.
- [34] Keane, Michael P. and Kenneth I. Wolpin, 2002, Estimating Welfare Effects Consistent with Forward-Looking Behavior, *Journal of Human Resources* 37.

- [35] Keilson, Julian and Adri Kester, 1977, Monotone Matrices and Monotone Markov Processes, *Stochastic Processes and Their Applications* (5), 231-241.
- [36] Kilian, Lutz, 2008, The Economics of Energy Price Shocks, *Journal of Economic Literature* (46), 871-909.
- [37] Lee, Sokbae, Oliver Linton, and Yoon-Jae Whang, 2009, Testing for Stochastic Monotonicity, *Econometrica* 77 (2), 585-602.
- [38] Lucas, Robert E., Jr., 1976, Econometric Policy Evaluation: A Critique, in K. Brunner and A. Meltzer Eds., *The Phillips Curve and Labor Markets*, Amsterdam: North Holland.
- [39] Mukherjee, Abhiroop, Manpreet Singh, and Alminas Zaldokas, 2017, Do Corporate Taxes Hinder Innovation?, *Journal of Financial Economics*.
- [40] Rosen, Harvey, 1992, *Public Finance*, Princeton University Press.
- [41] Rust, John, 2014, The Limits of Inference with Theory: A Review of Wolpin 2013, *Journal of Economic Literature*.
- [42] Schandlbauer, Alexander, 2017, How Do Financial Institutions React to a Tax Increase, *Journal of Financial Economics*.
- [43] Schepens, Glenn, 2016, Taxes and Bank Capital Structure, *Journal of Financial Economics*.
- [44] Slemrod, Joel, 1990, The Economic Impact of the Tax Reform Act of 1986, in J. Slemrod Ed., *Do Taxes Matter? The Impact of the Tax Reform Act of 1986*, MIT Press: Cambridge.
- [45] Summers, Lawrence, 1981, Taxation and Corporate Investment: A Q-Theory, *Brookings Papers on Economic Activity* 1, 65-105.
- [46] Tauchen, George, 1986, Finite State Approximations to Univariate Vector Autoregressions, *Economic Letters* 20 (2), 177-181.
- [47] Tauchen, George and Robert Hussey, 1991, Quadrature-Based Methods for Obtaining Solutions to Nonlinear Asset Pricing Models, *Econometrica* 59 (2), 371-396.
- [48] Wolpin, Kenneth, 2013. *The Limits of Inference without Theory*, MIT Press, Cambridge, MA.

Table I

Shock Responses in Economy with Decreasing Tax Rates

The table shows the model implied change in accumulation associated with a 5 percentage point change in the tax rate to one of its nearest neighboring tax rates. The assumed probability of a 5 percentage point tax rate decrease is 15% and the assumed probability of a 5 percentage point tax rate increase is 5%. The assumed risk free rate is 2.5% and the assumed depreciation rate is 7.25%.

Initial Tax Rate	Transition Down	Transition Up
0%	NA	-0.236
5%	0.236	-0.364
10%	0.364	-0.432
15%	0.432	-0.470
20%	0.470	-0.489
25%	0.489	-0.500
30%	0.500	-0.503
35%	0.503	-0.493
40%	0.493	-0.419
45%	0.419	NA

Table II

Shock Responses in Economy with Mean-Reverting Tax Rates

The table shows the model implied change in accumulation associated with a 5 percentage point change in the tax rate to one of its nearest neighboring tax rates. For tax rates no greater than 20%, the probability of a tax rate increase is 15% and the probability of a tax rate decrease is 5%. For tax rates greater than 20%, the probability of a tax rate increase is 5% and the probability of a tax rate decrease is 15%. The assumed risk free rate is 2.5% and the assumed depreciation rate is 7.25%.

Initial Tax Rate	Transition Down	Transition Up
0%	NA	-0.398
5%	0.398	-0.451
10%	0.451	-0.425
15%	0.425	-0.355
20%	0.355	-0.221
25%	0.221	-0.355
30%	0.355	-0.425
35%	0.425	-0.451
40%	0.451	-0.398
45%	0.398	NA

Table III

Estimated Tax Transition Rates

The table shows the discrete state space approximation for the estimated tax rate autoregressive process. The entries show the probability of a transition from a given rate into a given rate. Panel A assumes three possible tax rates and Panel B assumes four possible tax rates. The assumed risk free rate is 2.5% and the assumed depreciation rate is 7.25%.

Panel A		To		
		35.5%	50.4%	62.7%
From	35.5%	85.84%	14.13%	0.02%
	50.4%	12.24%	75.51%	12.24%
	62.7%	0.02%	14.13%	85.85%

Panel B		To			
		33.3%	45.9%	54.9%	67.4%
From	33.3%	83.0%	16.8%	0.2%	0.0%
	45.9%	13.5%	67.4%	19.0%	0.2%
	54.9%	0.2%	19.0%	67.4%	13.5%
	67.4%	0.0%	0.2%	16.8%	83.0%

Table IV

Ratio of Shock Response to Causal Effects

The table shows the model implied ratio of shock response to causal effect based upon the estimated tax transition rates shown in Table III. Panel A assumes three possible tax rates and Panel B assumes four possible tax rates. The assumed risk free rate is 2.5% and the assumed depreciation rate is 7.25%.

Panel A		To		
		35.5%	50.4%	62.7%
From	35.5%	NA	0.425	0.426
	50.4%	0.425	NA	0.427
	62.7%	0.426	0.427	NA

Panel B		To			
		33.3%	45.9%	54.9%	67.4%
From	33.3%	NA	0.440	0.485	0.478
	45.9%	0.440	NA	0.548	0.487
	54.9%	0.485	0.548	NA	0.443
	67.4%	0.478	0.487	0.443	NA

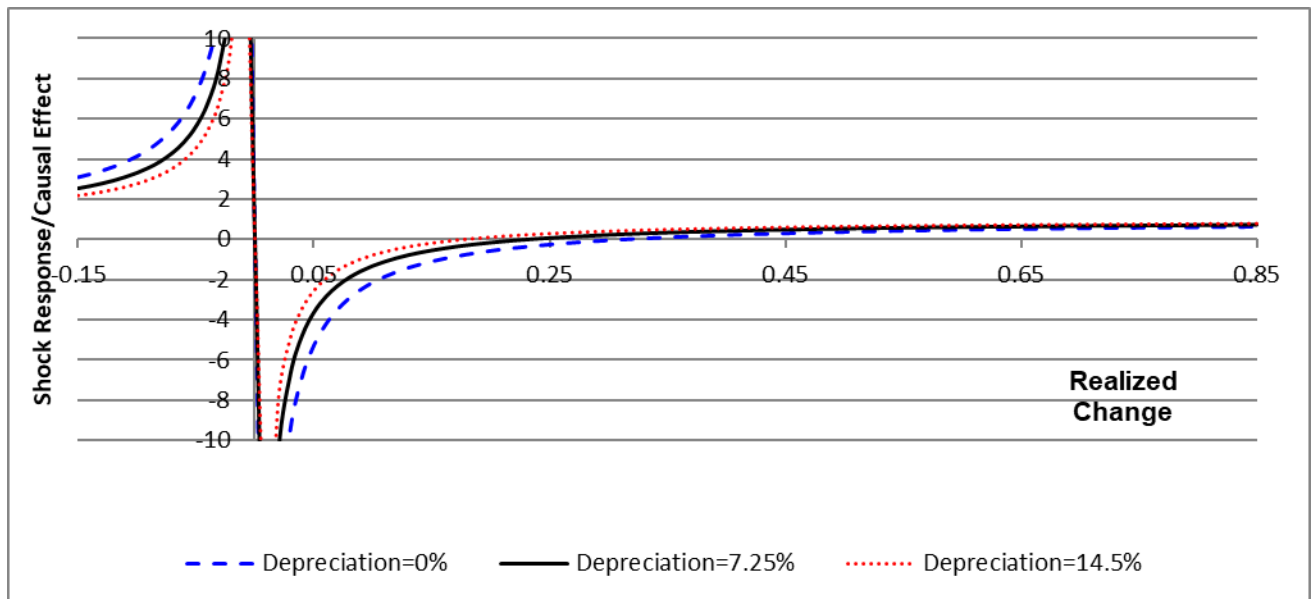


Figure 1. Ratio of Shock Reponse to Causal Effect. The figure shows the model implied ratio of shock response to causal effect with a 20 percent permanent shock probability drawn from a uniform distribution on the support from -0.15 to 0.85. The assumed risk free rate is 2.5 percent.