Estimating and Testing Dynamic Corporate Finance Models

Santiago Bazdresch
Banco de México

R. Jay Kahn
University of Michigan

Toni M. Whited
University of Michigan and NBER

We assess the finite sample performance of simulation estimators that are used to estimate the parameters of dynamic corporate finance models. We formulate an external validity specification test and propose a new set of statistical benchmarks that can be used to estimate and evaluate these models. These benchmarks are based on model policy functions. Our Monte Carlo simulations show that the estimators are largely unbiased with low root mean squared errors. When computed with an optimal weight matrix, the specification tests associated with the estimators are close to correctly sized. These tests have excellent power to detect misspecification. (*JEL* C14, C52, C61, G31, G32)

Received August 19, 2016; editorial decision May 30, 2017 by Editor Stijn Van Nieuwerburgh.

A large literature in finance and economics studies dynamic models of entrepreneurs, firms, and financial institutions, in which these agents, period by period, optimally make decisions about production, factor inputs, their compensation, and their financing.[1] Although these sophisticated dynamic programming problems are analytically complex and often only have approximate numerical solutions, this general research endeavor is promising.

This work would have been impossible without the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. This paper previously circulated under the title "Empirical Policy Function Benchmarks for Evaluation and Estimation of Dynamic Models." We have received valuable comments from the editor, Stijn Van Nieuwerburgh, two anonymous referees, Philip Bond, Gregor Matvos, Brent Glover, Missaka Warusawitharana, Andy Winton, and seminar participants at the Federal Reserve Board, the University of Minnesota, the University of Texas-Austin, the 2014 FIRS Meetings, the 2014 WFA Meetings, and the 2014 Summer Meetings of the Econometric Society. The opinions in this paper correspond to the authors and do not necessarily represent those of Banco de México. All code can be found on the *Review of Financial Studies* website. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Toni M. Whited, Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI 48109-1234; telephone: 734-764-1269. E-mail: twhited@umich.edu.

[1] For a comprehensive review of the corporate finance literature, see Strebulaev and Whited (2012).

doi:10.1093/rfs/hhx080 Advance Access publication July 14, 2017

Investment, labor demand, executive compensation, and financial decisions are intrinsically dynamic problems that can only have a quantitatively satisfactory representation in a dynamic model. Moreover, dynamic models allow researchers to extract a wealth of time-series and cross-sectional predictions with which to compare model and data. This richness allows researchers to discipline dynamic models more than static models. In general, this discipline is useful because it allows the evaluation of different models' ability to match the data and, ultimately, to establish quantitatively better theoretical bases for understanding firm behavior.

Despite the growing popularity of this research agenda, little work has been done to provide benchmarks and tests for assessing how these models fit the data. In this paper, we provide some initial inroads toward filling this gap, with an emphasis on three areas. We examine the finite sample properties of the simulated minimum distance estimators that have been used to estimate the parameters of dynamic models.[2] Second, we formulate an external validity test for these models. Third, we propose an alternative set of statistical benchmarks that can be used in the estimation and evaluation of dynamic models.

Our main analysis centers around a set of Monte Carlo experiments designed to evaluate the performance of simulated minimum distance estimators in a panel setting. On an intuitive level, these estimators work by choosing parameters that set moments (or functions of moments) computed from real data as close as possible to those computed from data simulated from a model. Examining the finite-sample properties of these estimators is potentially interesting and important because the estimators are closely related to closed-form generalized method of moments (GMM) estimators. It is well known that the finite-sample properties of GMM estimators can deviate from the asymptotic properties, even when the moment conditions have closed-form solutions (e.g., Hansen, Heaton, and Yaron 1996; Erickson and Whited 2002). In contrast, only very limited work has been done to understand the finite sample properties of simulation estimators (Michaelides and Ng 2000; Eisenhauer, Heckman, and Mosso 2015), largely because of the computational issues accompanying a Monte Carlo evaluation of an estimator that itself requires days of computation.

Advances in computing technology have allowed us to surmount this difficulty and find four basic results. First, all variants of the simulated minimum distance estimators that we examine produce parameter estimates that are nearly unbiased in finite samples even when the samples are substantially smaller than those typically available to corporate finance researchers. Moreover, these estimators often have extremely low root mean squared errors. Second, the standard errors that accompany these parameters produce reliable inference

[2] Hennessy and Whited (2005) and Cooper and Haltiwanger (2006) are early examples. More recent examples include Taylor (2010), Schroth, Suarez, and Taylor (2014), Warusawitharna and Whited (2016), Li, Whited, and Wu (2016), and van Binsbergen and Opp (2016).

only in the case in which the estimator is based on an optimal, clustered weight matrix. Third, we assess various model specification tests, finding that they over-reject strongly when we use an identity weight matrix or when the sample is small. However, these tests only over-reject slightly when we consider an optimal weight matrix and a larger sample size. This last result stands in sharp contrast to the documented strong over-rejection of GMM-based specification tests in some asset pricing contexts (e.g., Hansen, Heaton, and Yaron 1996). Fourth, the specification tests we consider also have excellent power to detect even small amounts of misspecification. Our result differs from the conclusion in Arellano and Bond (1991) that many panel GMM specification tests have poor power, but our result is in accord with the finding in Erickson and Whited (2012) that GMM specification tests can have excellent power to detect misspecification. We explain that these differences and similarities likely stem from the construction of the weight matrix, which, in our context, does not depend on parameter estimates.

Our second contribution is to formulate and evaluate the finite sample performance of tests to compare models and to assess the external validity of models. The model comparison tests we consider are from Nikolov and Whited (2014), who derive Wald tests to compare the equality of moments from different models. Our own external validity test statistics are a unique contribution of our paper. We derive tests of the null hypothesis of the equality of data and model-implied moments (or functions of moments), where these moments are not used in the estimation of the model. This type of test is useful for two reasons. First, it holds the model to a higher standard than a simple test of overidentifying restriction and thus accomplishes a purpose similar to that of an out-of-sample test. Just as an out-of-sample test assesses the ability of a statistical model to fit data not used in its estimation, our test assesses the ability of an economic model to fit economic predictions not used in its estimation. Second, such a test is useful for the simple reason that any model will fail if confronted with enough features of data that are generated by a world far more complex than the model. However, a test that can distinguish between features of the data that fit the model and those that do not can be of great use in directing researchers to develop better models.

Our third contribution is to provide guidance on the choice of which empirical predictions to use to evaluate, discipline, and test these models. Our motivation is twofold. First, we show with an example that benchmarks matter, as different benchmarks can produce economically distinct parameter estimates. Our second, broader motivation comes from the observation that different researchers make different arbitrary choices about which features of the data to consider. The result is a wealth of studies claiming that their models successfully explain the data. While the choice of empirical predictions naturally depends on the research question at hand and thus varies from application to application, there is still room for standardization of this choice. We argue that for any model there is a natural, intuitive set of statistics to be used for estimation and evaluation. Moreover, these statistics are comparable across different models. Therefore, they can be thought of as benchmarks that dynamic models should aim to match. In this sense, we provide one practical method to address the classic question in Gallant and Tauchen (1996) of which moments to match.

We derive these benchmarks from the model's policy functions, which characterize the solution of the model by stating the optimal response of the firm to its environment. Policy functions are thus the main objects that translate the assumptions of the model into a functional prediction about the firm's actions in different situations. Therefore, a direct, simple, and theoretically motivated way to evaluate a dynamic model is to evaluate its ability to replicate firms' observed policies. One way to accomplish this goal is to characterize firms' policy functions empirically, and then use these characterizations as the inputs for structural model estimation and evaluation.

Of course, for any particular model, there might be a unique feature of the data (such as the mean of a variable) that could be employed to estimate the model's parameters. As such, we do not argue that policy functions are the only data features useful for estimating models. Instead, we argue that empirical policy functions (EPFs) are common benchmarks that can be used as a starting point for model evaluation.

Using EPFs as benchmarks for estimating and evaluating models confers several advantages over the common practice of using arbitrary moments. First, the moments used in the traditional estimation of these models are computed by simulating data from the policy functions, so the policy functions in principle contain at least as much information as moments. The second argument is that these quantities are often already used in other types of structural estimation (Bajari, Benkard, and Levin 2007), albeit in sharply different ways. Thus, the use of policy functions as an input to the estimation of dynamic models is not a large departure from tradition. Third, applied researchers already estimate policy functions. For example, one of the most commonly estimated regressions in corporate finance is a regression of investment on cash flow (Fazzari, Hubbard, and Petersen 1988). However, this regression essentially constitutes a policy function from a model in which the firm observes its cash flow and then makes its investment decision. Thus, using policy functions as benchmarks to estimate the parameters of dynamic models affords a close connection between commonly run regressions and the dynamic models that underlie these regressions.

Although our paper is clearly related to the many applied papers that have used simulation estimators to estimate the parameters of dynamic models, it is also related to a set of applied econometrics papers in corporate finance that have sought to provide guidance to empirical researchers. For example, Petersen (2009) deals with the computation of standard errors in panel data, Erickson and Whited (2012) compare the finite-sample performance of estimators that treat measurement error, and Gormley and Matsa (2014) consider the treatment of unobserved heterogeneity. However, these papers deal with regression-based statistical analysis, while we look at methods used to estimate the parameters of dynamic economic models.

Next, our work stands apart from two more closely related papers. The first is the estimation method proposed by Bajari, Benkard, and Levin (2007). Their method uses estimated policy functions as direct inputs into an approximation of the model solution, which is subsequently used to construct moment inequalities that define the estimation. In contrast, our methods use a full model solution to generate simulated data. The empirical policy functions estimated on simulated data are then matched as closely as possible to policy functions estimated on real data. Second, our work is related to Gala and Gomes (2016). They focus on the estimation of the policy function of an investment model as a substitute for traditional regressions of investment on q. However, their work also contains an applied example of the idea of using policy functions as an input into indirect inference, where their application is the estimation of some parameters of their investment model.

Finally, although our methods are directly applicable to the estimation of the parameters of dynamic models, our introduction of empirical policy functions as benchmarks is also of use to the dynamic asset pricing literature, which often studies model calibrations. Even here a standardized benchmark could be of use in assessing the fit of various competing models, even though calibration offers no formal inference.

1. Trade-Off Model

This section outlines a class of simple dynamic capital structure models. As in any analysis of the finite sample properties of estimators, we need to choose a basic estimating equation. In the case of structural estimation, this choice is more involved than picking coefficients in a linear regression, as the estimating equation is a model itself. Thus, our goal is to pick a simple model that is as generic as possible.

The models of the firm that we consider are single-agent dynamic decision problems. In these models, the firm typically chooses optimal policies to maximize the expected discounted value of payout to current owners. Thus, the value of the equity of the firm can be described generically in terms of a Bellman equation:

$$\Pi(\boldsymbol{y},z)=\max_{\boldsymbol{y}'}\left\{e(\boldsymbol{y},\boldsymbol{y}',z)+\beta\mathbb{E}\Pi(\boldsymbol{y}',z')\right\}. \tag{1}$$

Here, $\boldsymbol{y}$ is a vector of endogenous state variables, z is a vector of stochastic exogenous state variables that follows a Markov process, $e(\boldsymbol{y},\boldsymbol{y}',z)$ is the current period net cash flow accruing to shareholders, $\Pi(\boldsymbol{y},z)$ is firm equity value, $\beta\in(0,1)$ is a discount factor, and $\mathbb{E}$ is the expectations operator with respect to the transition function for z. A prime indicates the next period, while the absence of a prime indicates the current period.

Given a functional form for $e(\boldsymbol{y},\boldsymbol{y}',\boldsymbol{z})$ and distributional assumptions for $\boldsymbol{z}$, both of which meet standard regularity condition e.g., (Stokey, Lucas, and Prescott 1989), the solution to this model exists. This solution can be expressed in terms of the value function, $\Pi(\boldsymbol{y},\boldsymbol{z})$, and the policy function, $\boldsymbol{y}'=G(\boldsymbol{y},\boldsymbol{z})$, which describes the firm's optimal choice of $\boldsymbol{y}'$, given the current state, $(\boldsymbol{y},\boldsymbol{z})$. The policy function is thus given by:

$$G(\boldsymbol{y},\boldsymbol{z})=\underset{\boldsymbol{y}'}{\operatorname{argmax}}\left\{e(\boldsymbol{y},\boldsymbol{y}',\boldsymbol{z})+\beta\mathbb{E}\Pi(\boldsymbol{y}',\boldsymbol{z}')\right\}. \tag{2}$$

Although our methods apply to the estimation of the parameters of any such model, to make our setting more concrete, we consider a special case that can be described as a streamlined version of the model in Hennessy and Whited (2005). We simplify this setting substantially. Otherwise, computing a Monte Carlo simulation of an estimator based on such a model would be infeasible.

In this model, the firm uses capital in a constant-returns technology to generate operating income according to zK, where K is the capital stock, and z is a profitability shock. Thus, in terms of the notation in Equation (1), $K\equiv\boldsymbol{y}$ and $z\equiv\boldsymbol{z}$. The use of a constant-returns technology follows Warusawitharna and Whited (2016) and greatly simplifies computation.

The profitability shock, z, is lognormally distributed and follows the process given by:

$$\ln(z')=\mu+\rho\ln(z)+\sigma\varepsilon', \qquad \varepsilon'\sim\mathcal{N}(0,1). \tag{3}$$

Each period the firm chooses investment, I, which is defined by a standard capital stock accounting identity:

$$K'\equiv(1-\delta)K+I, \tag{4}$$

in which δ is the rate of capital depreciation. We normalize the price of capital goods to one and assume that the firm faces real frictions in the form of convex investment adjustment costs. The function describing these costs is given by:

$$c(I)=I+\gamma K\frac{1}{2}\left(\frac{I}{K}\right)^2, \tag{5}$$

in which the first term represents the purchase price of capital goods, and γ quantifies investment adjustment costs. In many models of this type, investing incurs costs (Abel and Eberly 1994) that are independent of the amount of investment. For simplicity, we initially assume away these frictions, but we revisit the issues with fixed costs below in Section 5.

The firm's cash flow, $E^*(K,P,K',P',z)$, is its operating income plus its net debt issuance, $P'-P$, minus its net expenditure on investment, $c(I)$, and minus its interest payments on debt, rP:

$$E^*(K,P,K',P',z)=zK-c(I)+P'-P(1+r), \tag{6}$$

in which r is the risk-free rate of interest. We assume that $r<1/\beta-1$ to reflect the interest tax deduction and thus the tax benefits of debt. Motivated by the dynamic contracting literature (Rampini and Viswanathan 2013), we assume this debt is secured by capital, that is, we allow a fraction, ξ, of the capital to be used as collateral. Because some capital might be intangible and therefore of little worth to a lender, $0 \leq \xi \leq 1$. The collateral constraint can thus be expressed as:

$$P' \leq \xi K'. \tag{7}$$

This formulation of the leverage decision abstracts from debt issuance costs, but later we examine the ability of our estimation method to detect misspecification by considering the possible existence of issuance costs.

Cash flows to shareholders, $E(K,P,K',P',z)$, are defined in terms of the firm's cash flows, $E^*(K,P,K',P',z)$. A positive firm cash flow is distributed to its stockholders, while a negative cash flow implies that the firm instead obtains funds from shareholders. In this case, the firm pays a linear cost, λ. Thus, shareholder cash flows are given by:

$$\begin{aligned} E^* \geq 0 \quad &\Rightarrow \quad E = E^* \\ E^* < 0 \quad &\Rightarrow \quad E = E^*(1+\lambda). \end{aligned} \tag{8}$$

Having defined cash flows, we can now state the firm's problem as a special case of Equation (1):

$$\Pi(K,P,z) = \max_{K',P'} \left\{ E(K,P,K',P',z) + \beta \mathbb{E} \Pi(K',P',z') \right\}, \tag{9}$$

subject to Equations (4) and (7). Given easily verifiable restrictions on the parameters, this model satisfies the conditions in Stokey, Lucas, and Prescott (1989) for the existence of a solution.

We now simplify the model by exploiting our assumption of constant returns to scale and redefining all of the quantities in the model as a fraction of the capital stock, K. This transformation eliminates capital as a state variable and greatly simplifies computation. Define the following scaled variables:

$$p \equiv \frac{P}{K}, e \equiv \frac{E}{K}, i \equiv \frac{I}{K}, \pi(p,z) \equiv \frac{\Pi(K,P,z)}{K}.$$

Then, by dividing all of the variables in Equation (9) by K, we obtain the following Bellman equation:

$$\pi(p,z) = \max_{p',i} \left\{ e(p,p',i,z) + \beta \mathbb{E} \pi(p',z')(1-\delta+i) \right\} \tag{10}$$

and the constraints become:

$$e(p,p',i,z) = z(1-\tau) - i - \frac{\gamma i^2}{2} - p(1+r(1-\tau)) + p'(1-\delta+i), \tag{11}$$

$$p \leq \xi. \tag{12}$$

1.1 Optimal policies

Although much less elaborate than the model in Hennessy and Whited (2005), our simple model conveys much of the same intuition. To illustrate, we examine the optimality conditions for investment and leverage. To obtain the first-order condition for optimal investment, we differentiate Equation (10) with respect to i:

$$(1+\mathbb{I}_e\lambda)(1+\gamma i - p') = \beta \mathbb{E}\pi\left(p', z'\right), \tag{13}$$

in which $\mathbb{I}_e$ is an indicator function that is one if the firm is issuing equity. Naturally, this first-order condition appears similar to that from a neoclassical q model. If the firm is not issuing equity, then the marginal cost of investment is $(1+\gamma i - p')$, but in those states of the world in which the firm is issuing equity, this marginal cost rises by a factor of $(1+\lambda)$. At an optimum, the marginal cost of investment less the proceeds from any debt issues must equal the right-hand side of Equation (13), which is the expected future equity value per unit of capital. Because of our constant returns to scale assumption, this quantity is equal to the marginal value of capital or marginal q.

We obtain the first-order condition for debt by differentiating Equation (10) with respect to p', as follows:

$$1 = -\beta \mathbb{E}\left(\pi_p(p', z')\right). \tag{14}$$

Next, we use the envelope condition to eliminate $\pi_p(p', z')$ from the problem. Let η be the Lagrange multiplier associated with the collateral constraint in Equation (12). Substituting in the envelope condition, $-\pi_p(p,z) = (1+r)(1+\mathbb{I}_e\lambda)+\eta$, and rearranging gives:

$$1 = \beta \mathbb{E}\left((1+r)(1+\mathbb{I}_e\lambda')+\eta'\right). \tag{15}$$

Because of the assumption $r < 1/\beta - 1$, in the absence of financial frictions in the form of equity issuance costs, the obvious optimal policy of the firm is to borrow up to the collateral constraint. However, if the firm expects to be issuing equity in the next period, then Equation (15) will not hold at this corner solution. Instead, as is standard in dynamic investment models, debt capacity has value because it confers financial flexibility.

We now elaborate on this discussion by plotting the actual model policy functions in Panels A–C of Figure 1. Each panel depicts next-period net debt/capital, investment/capital, and net distributions/capital as a function of the profitability shock, and each panel is drawn for a different level of current debt/assets, in particular, the 10th, 50th, and 90th percentiles of simulated debt to assets. Of course, the shape of the policy functions depends on the parameters, so these functions are based on the parameter estimates from the moment estimation described below.

Several features of the policy functions shown in Figure 1 are noteworthy. First, they are clearly nonlinear. For example, when the firm is currently at a medium level of net debt-to-capital, distributions are unresponsive to the

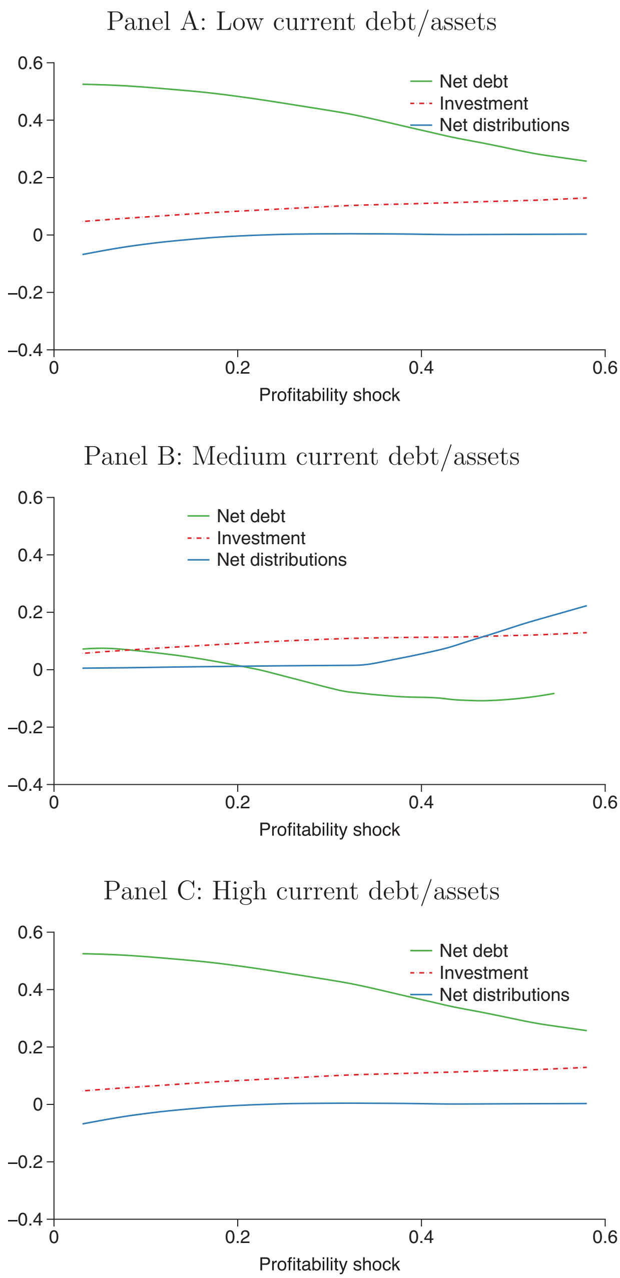

Figure 1
Theoretical policy functions
This figure plots optimal next-period net debt/capital, investment/capital, and net distributions/capital as a function of the profitability shock. Each panel is drawn for a different level of current debt/assets: low, medium, and high.

profitability shock if it is low, but they rise sharply with the shock if it is high. Second, while investment is always increasing in the profitability shock, the response is muted relative to the responses of debt and distributions. Quadratic investment adjustment costs lie behind this pattern, as they give the firm an incentive to smooth investment over time. Third, next-period net debt-to-capital increases with the profitability shock when current debt is low, but it decreases with the profitability shock when current debt is medium or high. This difference is due to the relative strengths of income and substitution effects. When debt is low and the firm has a great deal of free debt capacity, a substitution effect dominates. That is, when the firm receives a positive profitability shock, it optimally wants to invest in productive capital instead of debt capacity, so it funds its investment with debt, and both rise with the shock, z. However, when debt is at a medium or high level, an income effect dominates, so the firm optimally wants to invest in both capital and debt capacity, and we see a concurrent rise in investment and a fall in debt.

Our empirical policy function estimation benchmarks are based on estimates of these model policy functions instead of the actual policy functions themselves. Matching an exact function is ill-defined in a statistical sense, as indirect inference requires matching two sets of statistics: estimates from actual data with identical estimates from simulated data. It is this task to which we turn next.

2. Benchmarks and Estimation

2.1 Empirical policy functions

An empirical policy function is an estimate of the relationship between the current state of the firm and the policies the firm chooses in that state. One challenge posed by this estimation is the issue that in many models, some state and choice variables are unobservable. In such cases, it is often convenient to work with state and choice variables that are functions of the original variables. We therefore propose transforming the original policy function into one whose arguments are observable functions of the original arguments. Specifically, we define the transformed state and control variables as:

$$\boldsymbol{x} \equiv \boldsymbol{x}(\boldsymbol{y}, \boldsymbol{z}) \tag{16}$$

$$\boldsymbol{w} \equiv \boldsymbol{w}(\boldsymbol{y}'). \tag{17}$$

The dimensions of $\boldsymbol{x}$ and $\boldsymbol{w}$ are M and P, respectively. Note that these dimensions can differ from the dimensions of $(\boldsymbol{y}, \boldsymbol{z})$ and $\boldsymbol{y}'$. With these definitions, the policy function can be written as:

$$\boldsymbol{w} = H(\boldsymbol{x}). \tag{18}$$

For example, in the model in Section 1, z and p are the state variables, and i and p' are the control variables. The variables p, p', and i are obviously observable, and in the case of our simple constant returns model, z is the ratio of operating profits to capital. So $\boldsymbol{w} = \{i, p'\}$, and $\boldsymbol{x} = \{p, z\}$. In Section 5, we consider a decreasing returns-to-scale model in which we need to perform a nontrivial transformation.

With this notation in hand, we can explain the estimation of the key features of $\boldsymbol{w} = H(\boldsymbol{x})$. Linear regression is clearly inadequate for this task, as the constraints and nonconvexities in many dynamic models imply highly nonlinear policy functions, such as those in Figure 1. A natural alternative is any flexible semiparametric regression technique, as long as it can be characterized by a finite parameter vector. To describe the policy-function estimation step, we let $\boldsymbol{v}_{it} \equiv (\boldsymbol{w}_{it}, \boldsymbol{x}_{it}), i = 1, \ldots, N, t = 1, \ldots, T$, be a sample of observations on the state and control variables, where i indexes individual firms/plants/people, and t indexes time. For each control variable, we consider a semiparametric regression of the form:

$$\boldsymbol{w}_{it}^{n} = H^{n}(\boldsymbol{x}_{it}) + u_{it}^{n}, \tag{19}$$

in which the superscript n indicates the n^{th} element of the policy vector $\boldsymbol{w}_{it}$, and u_{it}^{n} is the regression disturbance, whose expectation, conditional on $\boldsymbol{x}_{it}$, is zero. This assumption is without loss of generality, as we do not need to find an interpretable causal relation between $\boldsymbol{w}_{it}$ and $\boldsymbol{x}_{it}$. Instead, estimation of the policy functions is motivated by a revealed-preference argument that the choices we observe in the data must be optimal choices, so the goal of the estimation is to uncover this endogenous, optimal relation between states and choices.

Because the policy function, $H(\boldsymbol{x}_{it})$, is of unknown form, to capture this nonlinearity we estimate Equation (19) using series approximating functions, which we denote as $h_j(\boldsymbol{x}_{it})$, $j = 1, \ldots, J$. We assume these approximate $H(\boldsymbol{x}_{it})$ in the following sense. As $J \to \infty$, the expected mean squared difference between $H(\boldsymbol{x}_{it})$ and a linear combination of the functions $h_j(\boldsymbol{x}_{it})$ approaches zero, that is:

$$\lim_{J \to \infty} E\left(\sum_{j=1}^{J} b_j h_j(\boldsymbol{x}_{it}) - H(\boldsymbol{x}_{it})\right)^2 = 0. \tag{20}$$

Several different series functions, such as power series or trigonometric series, can satisfy Equation (20), and we discuss the choice of the approximating function below.

2.2 Indirect inference

Once the above benchmarks are calculated, we use them to estimate a model through the indirect inference procedure in, for example, Smith (1993) and Gourieroux, Monfort, and Renault (1993). We now outline the procedure and then explain how it applies to our policy function benchmarks.

Recall that $\boldsymbol{v}_{it}$ is our vector of data observations. Let $\boldsymbol{v}_{it}^{s}$ be a simulated vector from simulation s, $s = 1, \ldots, S$, where S is the number of times the model is simulated. The simulated data vector, $\boldsymbol{v}_{it}^{s}(\theta)$, depends on a vector of structural parameters, θ. In our context, the structural parameters include the cost of equity issuance, λ, and the quadratic investment adjustment cost, γ. Next, we define the estimating equations as:

$$g(\boldsymbol{v}_{it},\theta)=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[m(\boldsymbol{v}_{it})-S^{-1}\sum_{s=1}^{S}m\left(\boldsymbol{v}_{it}^{s}(\theta)\right)\right], \tag{21}$$

in which $m(\cdot)$ is a vector of functions, whose dimension is at least as large as the dimension of the structural parameter vector, θ. For example, in the special case of a simulated moments estimator (Ingram and Lee 1991), $m(\cdot)$ is a vector of moments, but the indirect inference procedure is more general than a simulated moments estimator, so $m(\cdot)$ can also be a vector of functions of moments.[3] Hereafter, we refer to the vector $m(\cdot)$ generically as a benchmark. The objective of the estimation is to get this vector as close to zero as possible, so we introduce the term model error to describe the term in square brackets in Equation (21).

The indirect inference estimator, θ is the solution to the minimization of:

$$\hat{\theta}=\arg\min_{\theta} g(\boldsymbol{v}_{it},\theta)'\,\hat{W}_{nT}\,g(\boldsymbol{v}_{it},\theta), \tag{22}$$

in which $\hat{W}_{nT}$ is a positive definite matrix that converges in probability to a deterministic positive definite matrix W as $n\rightarrow\infty$.[4]

Finally, we explain how a simulated moments estimator and our use of empirical policy function benchmarks fit into this framework. In the case of a simulated moments estimation, the benchmark is simply a vector of moments. The case of our EPF benchmarks requires more explanation. First, we note that our implementation of indirect inference is closely related to the original motivation in Gourieroux, Monfort, and Renault (1993) for using this technique. They note that on an intuitive level, it might be desirable to estimate θ via maximum likelihood. However, in the class of computational dynamic models we consider, this strategy is unavailable, as the model solution does not provide an expression for the likelihood. Indirect inference fills this gap by using an auxiliary model, which should ideally capture important features of the data, even if it does not completely summarize the data in the same way that a likelihood function does. Because the empirical policy function is a characterization of the solution of the model, it is a highly suitable auxiliary model.

---

[3] Indirect inference also encompasses likelihood-based methods, which have not featured as prominently in problems of estimating the parameters of dynamic models, so we do not cover these methods here.

[4] While it is beyond the scope of this applied paper to state the asymptotic properties of these estimators, it is worth noting that, as is typical in panel settings, the asymptotic properties hold as n goes to infinity, with T held constant.

Because the auxiliary model is the empirical policy function given by Equation (20), the benchmark function $m(\boldsymbol{v}_{it},\theta)$ is an estimate of the parameter vector b in Equation (20). In both the real and simulated data, because we use ordinary least squares (OLS) to estimate b, $m(\cdot)$ is a vector of functions of moments, in particular, the means, variances, and covariances of the data. The simulated data depend on the structural parameter vector, θ, so the estimates of the auxiliary parameters in the simulated data, $b(\theta)$, naturally depend on the underlying structural parameters. The final goal is then to estimate θ by minimizing the distance between the parameter vector of the auxiliary model, b, estimated with a real data set and the same parameter vector estimated with data simulated from a model.

We next discuss the identification of the parameter vector, θ. Global identification requires a one-to-one mapping between the model parameters, θ, and a same-dimension subset of the parameters of the auxiliary model, b. Local identification of the parameters in any indirect inference estimation requires that the gradient of the auxiliary model with respect to the parameters, $\partial m\left(\boldsymbol{v}_{it}^{s}(\theta)\right)/\partial\theta$, have full rank. Intuitively, this condition indicates that for an element of the parameter vector, θ to be identified, some subset of the elements of $m(\cdot)$ must change when that particular element of θ moves. For example, identification of the depreciation rate, δ relies on the positive relation in the model between average investment and the depreciation rate. Although in the short run, investment is driven by productivity shocks, in the long run, the firm invests enough to replace depreciated capital, so high depreciation rates are associated with high average investment, and vice versa. In the EPF-based estimation, the intercept of the investment policy function serves the same purpose as the average depreciation rate.

Also important for identification and inference is the precision of the estimation of the auxiliary parameter vector, b, as the parameter vector θ is a function of b and thus, as shown in the expressions in the Appendix for the variance of θ, inherits its sampling variation. The consideration of precision brings up the important issue of the trade-off between using a close approximation to the true policy function and using up degrees of freedom in the data. Although it appears intuitive that a close approximation should provide sharper parameter identification, a better-fitting estimation of the policy function need not deliver better finite sample properties of an indirect inference estimator. The reason is sampling variation from the policy function estimation stage. For example, if one is estimating a model with a sharply discontinuous policy function, one needs a very flexible semiparametric estimator to capture the policy function shape. However, when such an estimator does deliver a good fit, it typically requires the estimation of so many auxiliary parameters that it ends up being high variance. This issue can compromise parameter identification because when the auxiliary parameters are estimated imprecisely, the amount of identifying information they provide is small.

It is beyond the scope and computational constraints of this paper to explore in detail the trade-off between sample size and the flexibility of the approximating function in Equation (20). Therefore, for simplicity, we consider a simple power series as our set of $h_j(\cdot)$ functions, with terms that are linear and quadratic in our state variables, as well as all cross products. As a robustness check, we also consider a higher order polynomial approximation.

Before continuing to explain inference and testing in this framework, it is worth digressing to explain the differences between our use of empirical policy functions in structural estimation and the method put forth by Bajari, Benkard, and Levin (2007) (BBL, hereafter). The BBL estimator is a two-step method that does not require a full model solution. The first step constitutes semi-parametric estimation of empirical policy functions, as well as transition probabilities for all stochastic state variables (z in our notation). These estimates are substituted directly into the model and are used to forward simulate value functions. The second step is the actual estimation of the structural parameters, which are chosen so that the observed policies are optimal relative to a set of perturbations of the value function.

This method has the distinct advantage of not requiring the time-consuming step of solving the model at every iteration of the econometric hill-descending routine. However, it does confer some disadvantages. First, the forward simulation step imposes a linear relation between the policy and value functions and thus introduces small sample bias, as value functions are generally nonlinear transformations of policy functions. For example, in a standard corporate finance setting, adjustment costs and equity issuance costs imply that leverage and investment do not enter linearly into the flow payoff of the firm. Our use of indirect inference sidesteps this issue because our policy function estimates are not introduced directly into the model solution. Second, BBL estimation requires that all state variables must be observed, as must all states to which a firm can transition with positive probability. This requirement is necessary for the forward simulation of the value function, yet it is unlikely to be satisfied in many corporate finance models, in which the underlying states are not observed (e.g., Hennessy and Whited 2005, 2007). Even when they are, such as in our constant-returns model, all of the state transitions may not be observable. For example, in the presence of model features such as default or equity issuance, the states that lead to either action can be reached from most other states with such low probability that, given typical sample sizes in Compustat, the state transitions may not be observed. The BBL estimator would struggle to detect the effect of these low probability events. In contrast, because policy functions reflect the existence of low-probability events even in states when they do not occur, the effect of these events can be seen in the comparison of model-implied and empirical policy functions. For example, a high cost of equity issuance affects the height and shape of the policy for optimal leverage over a wide range of states, not just the states in which equity issuance occurs.

2.3 Inference and tests

Much of our examination of the finite sample properties of these estimators centers around the performance of the test statistics that accompany indirect inference. As such, we now sketch the basic distributional properties of these estimators and then derive our external validity specification test, which is a new addition to this literature.

First, we tackle the question of the weight matrix. While any positive definite matrix, W, is a potentially valid choice, we focus on two typical choices: an identity matrix and the optimal weight matrix, which in a panel setting with large n and fixed T is the inverse of the clustered covariance matrix of the vector of functions, $m(\cdot)$. We denote this optimal weight matrix as $\hat{\Omega}^{-1}$ and calculate it using the influence function technique from Erickson and Whited (2002).

An influence function for an estimator is a function of the data whose mean has the same asymptotic distribution as the estimator. Thus, on an intuitive level, covarying the influence functions of two estimators produces their asymptotic covariance. Recall that the vector $m(\boldsymbol{v}_{it})$ contains moments (in the case of a simulated moments estimator) or functions of moments (in the case of indirect inference with EPF benchmarks). It follows that we can compute the clustered covariance matrix, $\hat{\Omega}$, by stacking the influence functions for the elements of $m(\boldsymbol{v}_{it})$ and simply taking a clustered covariance as follows:

$$\hat{\Omega} = \frac{1}{nT} \sum_{i=1}^{n} \left(\sum_{t=1}^{T} \psi_{m(\boldsymbol{v}_{it})} \right) \left(\sum_{t=1}^{T} \psi_{m(\boldsymbol{v}_{it})} \right)', \tag{23}$$

in which $\psi_{m(\boldsymbol{v}_{it})}$ is the vector of influence functions for the elements in $m(\boldsymbol{v}_{it})$. For example, in the case in which $m(\boldsymbol{v}_{it})$ is the vector of empirical policy function coefficients, b, the influence functions $\psi_{m(\boldsymbol{v}_{it})}$ are just standard OLS influence functions. Importantly, the expression given by Equation (23) does not depend on any model parameters, so the parameter vector never enters the weight matrix, as it does in many applications of GMM.

It is worth noting that other methods for calculating the covariance matrix $\hat{\Omega}$ are either cumbersome or potentially incorrect. For example, although estimation of $\hat{\Omega}$ as part of a large joint estimation of the vector $m(\boldsymbol{v}_{it})$ is asymptotically equivalent to our influence function approach (Taylor 2010), this type of exercise can be cumbersome if the dimension of $m(\boldsymbol{v}_{it})$ is large. Another possibility is to use a bootstrap. However, unless the resampling distribution is the same as the distribution of the data, this procedure is invalid, as the covariance matrix, $\hat{\Omega}$, is not an asymptotically pivotal statistic.

We leave the formulae for the standard indirect-inference test statistics to the Appendix, but we do sketch two tests. The first is the test from Nikolov and Whited (2014) of the null hypothesis that the simulated vector $m(\boldsymbol{v}_{it}, \theta_k)$ for a model k equals the vector $m(\boldsymbol{v}_{it}, \theta_j)$ for a different model j. A standard Wald test for this null hypothesis takes the form

$$\frac{nTS}{1+S}\left(m(\boldsymbol{v}_{it},\theta_k)-m(\boldsymbol{v}_{it},\theta_j)\right)'\left(\text{avar}\left(m(\boldsymbol{v}_{it},\theta_k)-m(\boldsymbol{v}_{it},\theta_j)\right)\right)^{-1}$$
$$\times\left(m(\boldsymbol{v}_{it},\theta_k)-m(\boldsymbol{v}_{it},\theta_j)\right),$$

in which $\text{avar}\left(m(\boldsymbol{v}_{it},\theta_k)-m(\boldsymbol{v}_{it},\theta_j)\right)$ is the asymptotic variance of the difference between the two moment vectors. As in the case of the estimation of the weight matrix given by Equation (23), we calculate this variance by covarying the influence function for $m(\boldsymbol{v}_{it},\theta_k)-m(\boldsymbol{v}_{it},\theta_j)$ with itself. Of course, this test can be performed for individual elements of $m(\cdot)$, subsets of elements, or for the entire vector. See Nikolov and Whited (2014) for details.

Second, we develop a Wald test for model errors that are not contained in the vector in Equation (21) used for estimation. Specifically, suppose we have a vector of benchmarks, $m^*(\cdot)$, that are not used to estimate the parameter θ. We want to test the null hypothesis that

$$g^*(\boldsymbol{v}_{it},\theta)=E\left(m^*(\boldsymbol{v}_{it})-S^{-1}\sum_{s=1}^{S}m^*\left(\boldsymbol{v}_{it}^{s}(\theta)\right)\right)=0. \tag{24}$$

This hypothesis constitutes a test of the external validity of the model, as it assesses the model's ability to explain patterns in the data that are not used to estimate its parameters. Under the null hypothesis that the model is correctly specified, this vector should equal zero. Because $g^*(\boldsymbol{v}_{it},\theta)$ is a function of the parameter vector θ, we can use a standard Wald test of the null in Equation (24). The hurdle is calculating the asymptotic variance of $g^*(x_i,b)$ because it is a function of two quantities that are estimated separately: the data vector, $g^*(\boldsymbol{v})$, and the parameter vector, θ.

Again, we use the influence function technique from Erickson and Whited (2002). The influence function for $g^*(\boldsymbol{v}_{it},\theta)$ can be calculated using the delta method, as follows. Let the influence function for observation it for θ be given by ϕ_θ, and let the influence function for $m^*(\boldsymbol{v}_{it})$ be ϕ_m^*. Then the influence function for $g^*(\boldsymbol{v}_{it},\theta)$ is given by:

$$\phi_g^*=\phi_m^*-\left(S^{-1}\sum_{s=1}^{S}\left(\partial m^*\left(\boldsymbol{v}_{it}^{s}(\theta)\right)/\partial\theta\right)\right)\phi_\theta.$$

The variance of $g^*(\boldsymbol{v}_{it},\theta)$ can then be obtained by covarying the influence function ϕ_g^* with itself, as follows:

$$\text{avar}(g^*(\boldsymbol{v}_{it},\theta))=E\left[\phi_g^*\phi_g^{*\prime}\right],$$

where the computation of this expectation proceeds as in Equation (23). The square-roots of the diagonal elements of $\text{avar}(g^*(\boldsymbol{v}_{it},\theta))$ can serve as standard errors in the construction of t-statistics of the null hypothesis that individual model errors in the vector $g^*(\boldsymbol{v}_{it},\theta)$ equal zero. Finally, the Wald test for the joint null hypothesis that all elements of the vector $g^*(\boldsymbol{v}_{it},\theta)=0$ can be constructed as

$$g^*(\boldsymbol{v}_{it},\theta)'(\text{avar}(g^*(\boldsymbol{v}_{it},\theta))^{-1}g^*(\boldsymbol{v}_{it},\theta), \tag{25}$$

which has degrees of freedom equal to the dimension of $g^*(\boldsymbol{v}_{it},\theta)$.

3. Estimation

3.1 Data

We draw our sample of firms from the Compustat database for the 1971 to 2015 period. We screen the sample as follows. The firm must have a CRSP share code of 10 or 11. We then drop all firms with fewer than two years of data or that belong to the financial (SIC code 6) or regulated (SIC code 49) sectors. We also drop quasi-governmental firms in SIC 9 and the U.S. Postal Service. Finally, we delete all observations in which any of the variables we use are missing, in which total assets are less than 10 million real 1982 dollars, or in which either sales or book assets grow by more than 200%. Even though we winsorize all variables at the 1% level, this last screen is nonetheless useful for minimizing the impact of outliers on our estimation. We are left with a sample of 111,902 firm/year observations.

We define the following variables to be used in the rest of the analysis. The numerator of book leverage is (DLC + DLTT − CHE) plus the capitalized value of leases, as in Li, Whited, and Wu (2016), and the denominator is total assets (AT). Profitability is OIBDP/AT, investment is CAPX/AT, and net payout is (CDIV + PDIV + PRSTK − SSTK)/AT. These data variables correspond to the variables p, z, i, and e in the model. Note that because the model variables p and e can be either positive or negative, they correspond to leverage net of cash and payout net of issuances. Finally, for reference, we also calculate the market-to-book ratio, which takes the form (AT + CSHO × PRCC_F − CEQ − TXDB)/AT.

The firms in Compustat are heterogeneous, yet the models described and exemplified above are typically models in which firms are ex ante homogenous. Thus, we demean each state and control variable at the firm level. This step is also important because we are concerned mainly with the dynamics of the different variables, as opposed to any cross-sectional variation. Nonetheless, we want the policy functions to be able to reconcile the average levels of the control variables in the model with the average levels in the data. Because removing firm fixed effects implies that all variables have a zero mean, we therefore add the mean over the entire sample back into each variable.

Table 1 presents summary statistics for the state and control variables in the model. Because we demean all variables at the firm level, the standard deviations and percentile ranges presented in Table 1 reflect within-firm variation, which is at most what a model of an individual firm can be expected

Table 1
Summary statistics: State and control variables

	Mean	SD	10%	Median	90%
Net book leverage	0.189	0.169	−0.009	0.190	0.383
Operating income/assets	0.142	0.090	0.044	0.141	0.243
Market-to-book ratio	1.587	0.700	0.954	1.523	2.252
Investment/assets	0.077	0.057	0.026	0.071	0.133
Net distributions/assets	0.001	0.076	−0.040	0.003	0.060
Observations	111,902				

This table provides summary statistics for the state and control variables in this paper. The sample is drawn from Compustat, covering an unbalanced panel of 111,902 firm/year observations from 1971 to 2015. All variables have been demeaned at the firm level, with the overall sample mean added back in, so standard deviations and percentile ranges reflect within-firm variation. Precise variable definitions are given in Section 3.1.

to explain. In this regard, the large ranges seen for the policy and state variables are of interest. For example, the firm at the 10th percentile of net leverage has negative net leverage, that is, it is actually holding more cash than debt. In contrast, the firm at the 90th percentile has net leverage of over 30%. The ranges for our other policy variables are similarly wide. In particular, the 10th and 90th percentiles of investment are consistent with the presence of some lumpiness in investment, with at least 10% of the firm-year observations exhibiting investment bursts that are nearly twice the median rate of investment. Finally, note that the standard deviation of the market-to-book ratio is much larger than the standard deviations of our state and policy variables. As much of this variation likely reflects measurement error (Erickson and Whited 2012), we do not use this variable in our structural estimations.

3.2 Estimation results

We now compare the results from estimating the trade-off model using two different types of estimation benchmarks: traditional moments and the series estimates of the empirical policy functions. Our intent is not to provide new economic insights, as the model we are estimating is only a simplified version of models that have been studied extensively (Hennessy and Whited 2005; DeAngelo, DeAngelo, and Whited 2011). Instead, our intent is to illustrate what one can learn from the differences in parameter estimates that different benchmarks produce. It is worth emphasizing that one should expect different parameter estimates when using different benchmarks, as models, which are simple by nature, can never reconcile all of the features of data that are generated by a complex world.

For the two estimation benchmarks we consider, we need to estimate seven parameters: δ, the depreciation rate of capital, λ, the equity issuance cost parameter, ξ, the collateral parameter, and γ, the adjustment cost parameter, as well as μ, ρ, and σ, the mean, serial correlation, and standard deviation of the driving process for $\ln(z)$.

For our first, moments-based estimation, we need at least seven moments to estimate these seven parameters. We choose the means and variances of the four variables in our model: investment, leverage, net equity payout, and operating profits. This choice produces an overidentified model by one degree of freedom.

For our second, EPF-based estimation, we use a second-order polynomial approximation to the policy functions. Specifically, we regress the three policy variables—net leverage, net distributions to shareholders, and investment—on second-order polynomials in the two lagged state variables—net leverage and operating profits. Thus, each regression contains six parameters: an intercept, the two coefficients on the linear terms, the two coefficients on the quadratic terms, and the coefficient on the interaction term. With three empirical policy functions, we end up with an $m(\cdot)$ vector that contains 18 elements, which overidentifies the model by 11 degrees of freedom. In both the moments-based and the EPF-based estimation, the weight matrix for measuring the distance between the simulated and data moments is the inverse of the clustered covariance matrix of either the moments or the OLS policy function coefficients.

Panel A of Table 2 presents the parameter estimates from the two estimations. Several general patterns emerge. First, with the exception of the equity issuance cost parameter, λ, all of the coefficients are statistically significantly different from one another across the two estimations, and many pairs of coefficients are also economically different. For example, the collateral parameter, ξ, is 50% higher in the EPF-based estimation, and the shock variance, σ, is almost 30% higher in the EPF-based estimation. These sharp differences are to be expected, as our stylized model is much simpler than the typical model used in the literature.

One result in particular highlights the usefulness of the two different benchmarks for identifying parameters. The large standard error on the estimate of the equity issuance cost, λ, in the moments-based estimation means that few of the moments change significantly with λ. Indeed, only the mean and variance of net distributions change with this parameter, and the effects are small. In general, poor parameter identification results in larger standard errors. In turn, larger standard errors imply that parameter values far from the point estimates produce largely the same set of estimated benchmarks as the actual point estimates themselves.

In Panel A of Table 2, we also present results from two specification tests: the standard test of overidentifying restrictions and the external validity test given by Equation (25). In the case of the moments-based estimator, the moments used for the external validity test are the empirical policy function coefficients, and in the case of the EPF-based estimator, the moments used for the external validity test are the moments used in the moments-based estimator. Both tests show that the model is strongly rejected. Again, this result is to be expected, as we are dealing with an oversimplified and thus naturally misspecified model. This result is also important, as it foreshadows the result from our Monte Carlo exercises that these two specification tests have excellent power to detect even small amounts of model misspecification.

Table 2
Estimations of trade-off model with different benchmarks

Panel A: Parameters

Parameter	Moments-based	EPF-based
μ	−2.2067	−2.0455
	(0.0236)	(0.0019)
ρ	0.8349	0.8200
	(0.0003)	(0.0013)
σ	0.3594	0.4497
	(0.0211)	(0.0012)
δ	0.0449	0.0665
	(0.0011)	(0.0001)
γ	29.9661	21.5488
	(2.6340)	(0.2225)
ξ	0.3816	0.6181
	(0.0038)	(0.0043)
λ	0.1829	0.2350
	(1.2196)	(0.0282)
Overidentifying χ^2 p-value (d.f.)	0.000 (1)	0.000 (11)
External validity χ^2 p-value (d.f.)	0.000 (18)	0.000 (8)

Panel B: Moments

	Data	Moments-based		EPF-based	
		Simulated	t-statistic	Simulated	t-statistic
Mean leverage	0.1886	0.1774	6.5603	0.1984	−2.8716
Variance leverage	0.0285	0.0225	33.6700	0.1110	−68.6243
Mean investment	0.0768	0.0600	60.6440	0.0809	−6.0456
Variance investment	0.0033	0.0002	67.1409	0.0007	41.2507
Mean distributions	0.0012	0.0124	−107.1251	0.0020	−1.0780
Variance distributions	0.0058	0.0026	86.1149	0.0010	45.5955
Mean profits	0.1421	0.1325	285.1138	0.1689	−24.3985
Variance profits	0.0081	0.0070	22.3168	0.0170	−38.3895

Panel A contains parameter estimates from a trade-off model using two different benchmarks: traditional moments and empirical policy functions (EPFs). Indirect inference is performed by minimizing the (inverse covariance matrix weighted) distance between the simulated values of each set of benchmarks and the corresponding values found in Compustat. Clustered standard errors are in parentheses under the parameter estimates, and degrees of freedom are in parentheses next to the specification test statistic p-values. μ is the (log) mean of the productivity process, and ρ and σ are the serial correlation and residual standard deviation of this process. δ is the depreciation rate of capital, γ is the investment quadratic adjustment cost, ξ is the collateral value of capital, and λ is the linear equity issuance cost. In Panel B, the first column presents the data estimates used in the traditional moments-based estimation. The second and third columns contain the simulated moments from the traditional moments-based estimation, along with the t-statistics for the difference between the actual and simulated moments. The next two pairs of columns contain analogous results for the two EPF-based estimations.

In Panel B of Table 2, we compare the model-implied simulated moments from the moments-based and EPF-based estimations. The first column contains the data moments used in the traditional moments-based estimation. The second and third columns contain the simulated moments from the traditional moments-based estimation and the t-statistics for the difference between the data and simulated moments—that is, the t-statistic for the model error. The next two columns contain analogous estimates and test statistics for our EPF-based estimation, so these two columns represent an external validity test of the model in the EPF-based estimation.

In Panel B of Table 2, we find large t-statistics on nearly all of the moment conditions. These large figures stem from two sources. The first is the highly stylized nature of this particular model. Second, because the t-statistics are proportional to the square root of the sample size, a large sample size of over 100,000 implies precisely estimated moments. Next, we usually find smaller t-statistics on the external validity moment conditions for the EPF-based estimation than on the moment conditions that are actually used for the moments-based estimation. Intuitively, one would expect external validity tests to reject more strongly, so this result at first glance seems odd. However, the result can be explained by the observation that the estimated parameters and the simulated moments inherit the sampling variability of the benchmarks used to estimate them, whether the benchmarks are moments or estimates of empirical policy functions. Because regression coefficients are typically estimated with less precision than means in an i.i.d. cross section, the simulated moments from the EPF-based estimation have more sampling variability and consequently are associated with lower t-statistics.[5]

4. Monte Carlo Simulations

Although indirect inference estimators are asymptotically consistent and can be efficient within a class of minimum distance estimators, little is known about their finite sample properties, which need not mirror their asymptotic properties. This section describes a set of Monte Carlo experiments that assess the finite sample performance of the indirect inference estimators that we have discussed and used so far. We design these experiments as follows. Each Monte Carlo is based on 1,000 simulated data sets,[6] in which the data are simulated from the actual (not estimated) policy functions that characterize the solution to the model in Section 1. We consider two sample sizes for our simulated data. The first has a length of 8 and a cross-sectional width of 9,375, for a total size of 75,000. These dimensions are slightly smaller than those of our actual data. We also consider a small sample size, with a length of 8 and a cross-sectional width of 125, with a total size of 1,000.

We create our simulated samples as follows. First, we choose values for four key parameters: the depreciation rate, δ, the cost of equity issuance, λ, the collateral parameter, ξ, and the convex investment adjustment cost parameter, γ. To make the Monte Carlo simulations relevant to our application, we set these parameters equal to the estimates from the EPF-based estimation in Table 2. Because estimations with more parameters take more time, to keep the Monte Carlo tractable, we treat the parameters that govern the process for z (μ, σ, and ρ) as known, again setting them to the estimates from the EPF-based estimation in Table 2. We then solve the model to obtain the policy functions and simulate

5 We have also done a symmetric analysis with the coefficients of the estimated policy functions. Because the results are largely similar, we omit them.

6 Because one estimation of a model takes several hours, larger numbers of Monte Carlo trials, which are typical in many simulation studies, are infeasible.

1,000 different data sets from the policy functions. Most of our analysis is based on simulations in which we estimate the same model that generated the data, so the null hypothesis underlying these experiments is that the model is true. Later we examine the performance of the estimators when the model used to simulate the data differs from the model used for estimation.

We consider two weight matrices: the optimal clustered weight matrix and an identity weight matrix, which is popular in the macroeconomics and asset pricing literatures that focus on the calibration of models. For each of these weight matrices, we then estimate the model using both traditional moments and empirical policy functions as benchmarks.

From an ex ante perspective, the choice of weight matrix is not obvious, as neither weight matrix necessarily puts the most weight on features of the data that might be of the most economic interest in any particular application. For example, although the optimal weight matrix minimizes overall model error by (roughly) putting the most weight on the most precisely estimated moments, these particular moments need not be the most relevant to the economic questions addressed by the model. The use of identity weight matrices confers a different disadvantage, as identity weight matrices have the unfortunate property that they mechanically put the most weight on the moment that is largest in absolute value. Similarly, minimizing the percent difference between the moments mechanically puts the most weight on the smallest moment in absolute value. It is hard to imagine an economic objective that coincides with either of these mechanical objectives. This perspective lies in contrast to the advice given in Cochrane (2005) regarding the use of an identity matrix in tests of asset pricing models. However, this advice makes sense in the context of minimizing pricing errors because the returns on different portfolios are typically all of the same order of magnitude. In contrast, moments used in a corporate finance simulated moments exercise can be of very different magnitudes, as can the regression coefficients that define an empirical policy function, so an identity matrix can end up mechanically emphasizing uninteresting moments. In the end, because there is no obvious ex ante choice, the consideration of the finite sample properties of these two weight matrices is of particular interest.

4.1 Baseline results

Table 3 shows the results from our first Monte Carlo simulation, where we set the sample size to 75,000. For each parameter we estimate, and for each method we use, we report the mean bias, the root mean squared error (RMSE), and the probability of rejecting a nominal 5% test of the null hypothesis that the parameter equals its true value. Bias and RMSE are both expressed as a percentage of the true parameter.

Three notable results stand out. First, both the moments-based and EPF-based estimators produce largely unbiased parameter estimates, regardless of the weight matrix we use. Second, except for the case of the equity issuance

Table 3
Monte Carlo comparison of simulation estimators: Large sample size

	Moments-based		EPF-based	
Parameter	Identity	Clustered	Identity	Clustered
δ (depreciation rate)				
Average % bias	0.123	−0.006	−0.021	−0.001
RMSE %	0.608	0.047	0.059	0.010
Pr(t)	0.605	0.367	0.309	0.348
λ (equity issuance cost)				
Average % bias	0.598	−0.165	1.793	−0.047
RMSE %	3.141	1.477	2.662	0.875
Pr(t)	0.001	0.002	0.001	0.003
ξ (collateral parameter)				
Average % bias	−0.189	−0.299	−0.308	0.035
RMSE %	0.790	0.659	1.997	0.110
Pr(t)	0.117	0.277	0.359	0.019
γ (investment adjustment cost)				
Average % bias	−0.239	0.027	0.052	0.007
RMSE %	1.273	0.106	0.127	0.022
Pr(t)	0.313	0.135	0.115	0.100

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1. We consider two different benchmarks: traditional moments and empirical policy functions. Both estimators use a clustered weight matrix. For each parameter, we report three statistics. Bias is expressed as a fraction of the true coefficient value. RMSE indicates root mean squared error and is also expressed as a fraction of the true coefficient. Pr(t) is the fraction of the time we observe a nominal 5% rejection of the null hypothesis that a parameter equals its true value using a t-test.

cost, λ, all estimators have extremely low RMSEs. Even the slightly higher RMSEs for λ are not large. For example, the RMSE for the moments-based estimator with the identity weight matrix is only 3.14% of the true parameter value of 0.235. Third, using a clustered weight matrix produces estimators with lower bias and RMSE for both the moments-based and EPF-based estimators.

These results are interesting in that they are the only extant evidence on the finite sample performance of the sorts of simulation estimators used in corporate finance. And the evidence is encouraging. If the model is correctly specified, as it is here, then the estimates from these simulation methods have both low finite-sample bias and variance. Intuitively, this good performance stems from two sources. First, the parameter estimates inherit the sampling variation of the moments and OLS regression coefficients used in the estimation, and these simple benchmark statistics are themselves precisely estimated in finite samples. Second, the mappings from the moments to the parameters are in most cases steep, with the one exception again being the estimators for λ. Equation (A1) in the Appendix shows that this weak mapping should be evident in the parameter standard error, which, for λ, exceeds the parameter value itself on average (not reported).

The evidence on the finite sample performance of the t-tests associated with each of these parameters is somewhat less encouraging. First, only in a few instances do we find that the actual rejection rates of these tests approach the nominal 5% values. Second, for the parameter λ, these tests almost never produce a rejection, again because of the large standard errors associated with

Table 4
Monte Carlo comparison of simulation estimators: Small sample size

Parameter	Moments-based Identity	Moments-based Clustered	EPF-based Identity	EPF-based Clustered
δ (depreciation rate)				
Average % bias	0.848	0.183	0.055	−0.005
RMSE %	1.685	0.360	0.230	0.099
Pr(t)	0.891	0.487	0.344	0.501
λ (equity issuance cost)				
Average % bias	−0.851	0.779	−2.715	0.504
RMSE %	6.414	3.623	6.878	3.046
Pr(t)	0.019	0.000	0.001	0.001
ξ (collateral parameter)				
Average % bias	−1.841	−0.027	−3.065	−0.060
RMSE %	4.220	2.843	5.917	1.867
Pr(t)	0.087	0.042	0.317	0.191
γ (investment adjustment cost)				
Average % bias	−1.719	−0.381	−0.093	0.023
RMSE %	3.443	0.761	0.482	0.208
Pr(t)	0.834	0.203	0.159	0.180

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of size 1,000. See Table 3 for definitions.

this parameter. Third, the tests associated with the identity weight matrices tend to reject the null much more often than the tests associated with the clustered weight matrix. For example, for the EPF-based estimator of the collateral parameter, ξ, the t-test rejection rate is 0.019 when we use the optimal weight matrix, but the rejection rate rises to 0.359 when we use an identity weight matrix. We conjecture that the often higher rejection rates for the tests associated with the identity weight matrix are due to the inevitable inaccuracy in the numerical gradients used to calculate the parameter standard errors. Inspection of the indirect inference variance formula given by Equation (A1) in the Appendix reveals that small errors in the gradient can be compounded to reduce the estimator variance. In contrast, the estimator variance expression in Equation (A2) in the Appendix is less likely to suffer from this problem, as it derives from the optimal weight matrix

Next we turn to the results in Table 4 from a similar Monte Carlo simulation in which the sample size is set at 1,000. In this case, we again find that the estimates of all of the parameters are nearly unbiased, and that the estimators that use the optimal weight matrix outperform those that use the identity weight matrix. However, the substantially smaller sample size results in much higher RMSEs.

From comparing Tables 3 and 4, we can conclude that the indirect inference estimators used in the estimation of dynamic models produce highly accurate parameter estimates when the model is correctly specified. We can also conclude that inference about the parameter estimates is much less accurate, especially in small samples, and especially in the case of a diagonal weight matrix. The high rates of test rejection imply that the t-statistics associated with the parameter

Table 5
Monte Carlo comparison of specification tests

	Moments-based		EPF-based	
Parameter	Identity	Clustered	Identity	Clustered
Sample size = 75,000				
Overidentification test rejection rate	0.558	0.048	0.825	0.083
External validity test rejection rate	0.985	0.843	0.668	0.079
Moment *t*-statistics:				
maximum rejection rate	0.317	0.022	0.354	0.024
median rejection rate	0.132	0.012	0.056	0.010
minimum rejection rate	0.000	0.005	0.000	0.005
Sample size = 1,000				
Overidentification test rejection rate	0.470	0.081	0.719	0.116
External validity test rejection rate	0.973	0.713	0.607	0.417
Moment *t*-statistics:				
maximum rejection rate	0.288	0.049	0.397	0.047
median rejection rate	0.066	0.026	0.116	0.000
minimum rejection rate	0.000	0.014	0.000	0.000

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of sizes 75,000 and 1,000. The samples are generated from the model in Section 1. The moments-based estimator minimizes the distance between simulated and data moments. The EPF-based estimator minimizes the distance between policy functions estimated from simulated data and those estimated from real data. We report the fraction of trials that produce a nominal 5% rejection of three additional tests. The first is the test of the model overidentifying restrictions. The second is our external validity test, which has two varieties. For the EPF-based estimator, it is a chi-squared test of the null hypothesis that the moments equal their true values. For the moments-based estimator, it is a chi-squared test of the null hypothesis that the policy function slopes equal their true values. The third is a *t*-test on individual moment conditions. For these tests, we report the highest, median, and lowest rejection rates.

estimates in Table 2 need to surpass critical values higher than the usual 1.96 threshold for a nominal 5% test.

Table 5 presents additional results from the two simulations, specifically, the performance of various specification tests associated with the moments-based and EPF-based estimators. We examine three tests, all of which are given in the Appendix. The first is the standard test of the overidentifying restrictions of the model, which is given by Equation (A4) in the case of the diagonal weight matrix or Equation (A5) in the case of the optimal weight matrix. The second is the external validity test given by Equation (25). As in the data analysis in Section 3, for the moments-based estimator, the external validity moments are the empirical policy function coefficients, and for the EPF-based estimator, the external validity moments are the moments used in the moments-based estimator. The third test is a *t*-test of the null hypothesis that an individual element of the benchmark vector $g(\boldsymbol{v}_{it},\theta)$ equals zero. For brevity, instead of reporting rejection rates for each element of $g(\boldsymbol{v}_{it},\theta)$, we report the minimal, median, and maximal rejection rates across all of the moments (or empirical policy function coefficients) used in the estimation.

The results in Table 5 show that the performance of all of these test statistics is much better when the estimation uses an optimal weight matrix. For both the large and small sample size, and for both the moments-based estimator and the EPF-based estimators, the external validity test and the overidentification test over-reject strongly when the estimation employs an identity weight matrix. This result is intuitive because estimations that use an identity weight matrix essentially put the most weight on benchmarks (moments or policy function coefficients) that are largest in absolute value. This objective clearly fails to coincide with the sensible objective of these tests, which is to detect overall model error, in which, roughly speaking, precisely estimated benchmarks are given more weight than imprecisely estimated benchmarks. Thus, we naturally observe large rejection rates when the weight matrix is an identity matrix.

In contrast, when we look at the estimators that use the clustered weight matrix, the actual size of most of the specification tests is within ten percentage points of the 5% nominal size. The one notable exception is the external validity test in the case of the moments-based estimator. This result makes intuitive sense. The intercept and slope coefficients from the empirical policy functions implicitly contain all of the information in the means and variances of the policy variables. However, they also contain information about covariances, which the moments-based estimator does not use. Thus, it is no surprise that the moments-based estimator cannot match the policy coefficients but that the converse holds—that is, the EPF-based estimation can match the moments.

Table 5 also shows the closeness of the nominal and actual sizes of most of the *t*-tests for the null hypotheses that the elements of $g(\boldsymbol{v}_{it},\theta)$ equal zero. This relatively good performance is evident even for the estimators that use an identity weight matrix.

Figures 2 and 3 add texture to the results in Table 5 by examining the distributions of three specification test statistics: the test of overidentifying restriction, the external validity test, and the test of overidentifying restrictions evaluated at the true parameters. Figure 2 shows the results from the simulation with the large sample size of 75,000. On the *y*-axis of each plot is the percentile over all Monte Carlo trials of each chi-squared statistic. On the *x*-axis is the theoretical percentile of that statistic, which has an asymptotic chi-squared distribution. If the Monte Carlo distribution of the chi-squared statistic equals its theoretical distribution, then the simulated and theoretical percentiles should plot along the 45-degree line. If the tests over-reject (under-reject), then the simulated and theoretical percentiles should plot below (above) the 45-degree line.

Panel A of Figure 2 considers the moments-based estimator. In the case of both the optimal weight matrix and the identity weight matrix, if we evaluate the model at the true parameter vector, the overidentification test statistic has a chi-squared distribution, as it should. This result is intuitive in that this Monte Carlo experiment isolates sampling variation. In the estimation with an optimal weight matrix, the test of overidentifying restrictions for the model evaluated at the estimated parameters is close to its theoretical distribution. However, the external validity test deviates sharply. This deviation is driven by a large fraction of extremely large test statistics, with 50% of the simulated statistics lying near the 100th percentile of the theoretical chi-squared distribution. In

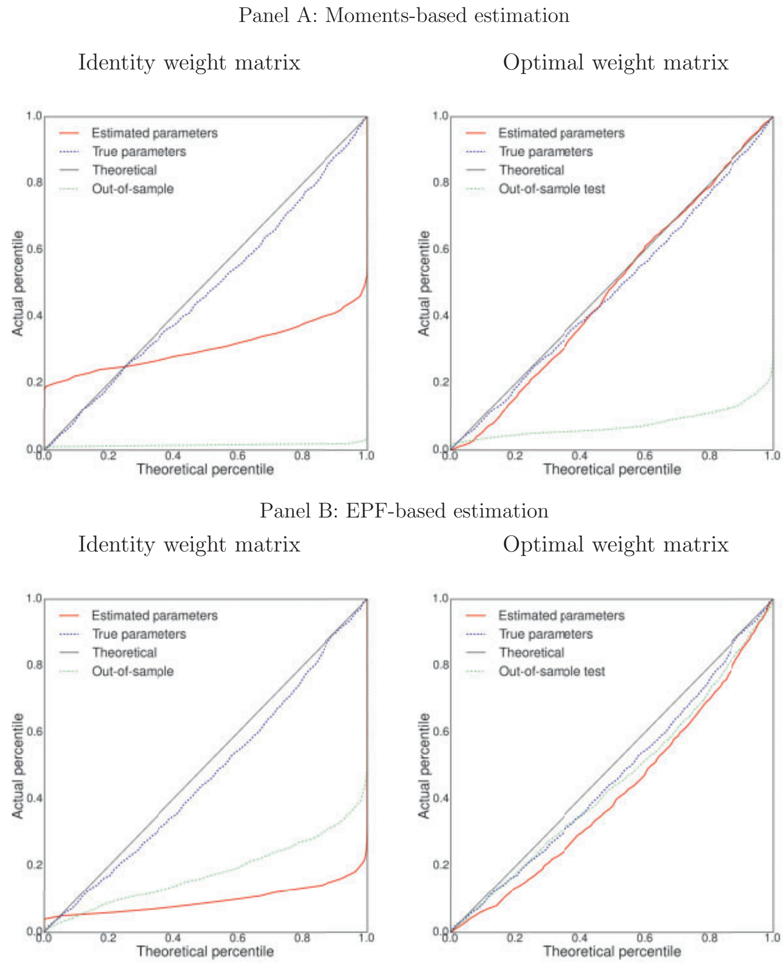

Figure 2
Distribution of chi-squared test statistics: Large sample size
Indicated probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1, and estimation uses both an identity and a clustered weight matrix. On the *y*-axis of each plot is the percentile over all Monte Carlo trials of each chi-squared statistic. On the *x*-axis is the theoretical percentile of that statistic. If the Monte Carlo distribution of the chi-squared statistic equals its theoretical distribution, then the simulated and theoretical percentiles should plot along the 45-degree line. We show three statistics. The first is a test of overidentifying restrictions when the model is evaluated at the true parameter vector. The second is the standard test of overidentifying restrictions using the estimated parameter vector, which we refer to as the minimizing parameters. The third has two varieties. For the moments-based estimator, it is a chi-squared test of the null hypothesis that the policy function slopes equal their true values. For the EPF-based estimator, it is a chi-squared test of the null hypothesis that the moments equal their true values.

the estimation with an identity weight matrix, both the overidentification test and the external validity tests exhibit deviations from the 45-degree line, with both tests over-rejecting strongly.

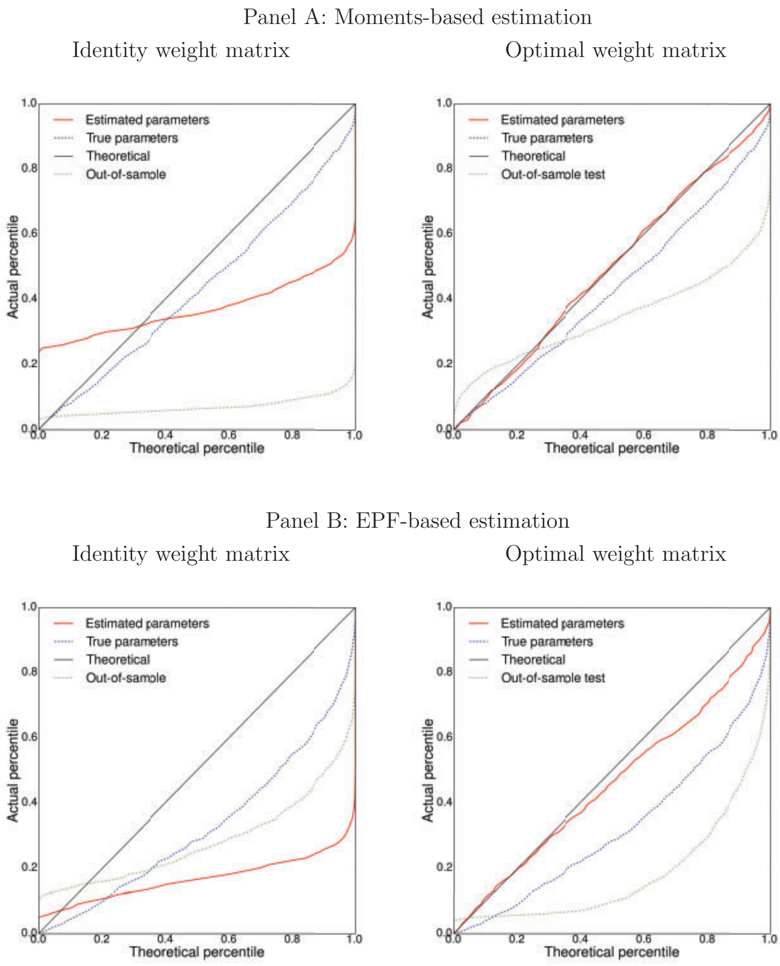

Figure 3
Distribution of chi-squared test statistics: Small sample size
Indicated probabilities are estimates based on 1,000 Monte Carlo samples of size 1,000. The samples are generated from the model in Section 1, and estimation uses both and identity and a clustered weight matrix. On the *y*-axis of each plot is the percentile over all Monte Carlo trials of each chi-squared statistic. On the *x*-axis is the theoretical percentile of that statistic. If the Monte Carlo distribution of the chi-squared statistic equals its theoretical distribution, then the simulated and theoretical percentiles should plot along the 45-degree line. We show three statistics. The first is a test of overidentifying restrictions when the model is evaluated at the true parameter vector. The second is the standard test of overidentifying restrictions using the estimated parameter vector, which we refer to as the minimizing parameters. The third has two varieties. For the moments-based estimator, it is a chi-squared test of the null hypothesis that the policy function slopes equal their true values. For the EPF-based estimator, it is a chi-squared test of the null hypothesis that the moments equal their true values.

Panel B presents the results for the EPF-based estimator. As expected, given the results in Table 5, all of the tests associated with the estimator that uses the optimal weight matrix plot near the 45-degree line. However, for the identity weight matrix, we again find a large fraction of the tests statistics near the one-hundredth percentile of the theoretical distribution.

Figure 3 presents analogous results for the simulation with a small sample size of 1,000. While these results largely mirror those in Figure 2, there is one important difference. Even when the model is evaluated at the true parameters, the actual percentile of the overidentification test statistic fails to mirror the theoretical percentile, and this pattern is more pronounced for the EPF-based estimator. Intuitively, sampling variation in the actual estimation of the data moments is more important with a small sample size. Moreover, estimating empirical policy functions requires more data than estimating moments, so the extra sampling variation adds noise to the specification test statistics.

In conclusion, the good performance of many of these tests appears unusual, given the well-documented tendency of tests associated with GMM estimators to over-reject in finite samples (Hansen, Heaton, and Yaron 1996; Shanken and Zhou 2007). We attribute the good performances of these estimators to two factors. The first is the large sample size, which is not a feature of many time-series asset pricing or macroeconomics applications. The second and more important factor is the weight matrix. Because the data and parameters are additively separable in Equation (21), it is possible to estimate the optimal weight matrix without knowledge of the parameter vector. This separability property is also a feature of a small number of GMM applications, including asset pricing applications based on minimizing pricing errors and the estimators in Erickson and Whited (2002). Interestingly, these latter estimators also have good finite sample properties.

4.2 Test power

In all of the simulations discussed above, the model we estimate is the same as the one that generates the data. In other words, we have examined the properties of the simulation estimators we consider under the maintained hypothesis that the model is correct. However, models are by nature misspecified. Therefore, we next examine whether the specification tests that accompany simulation estimators have power to detect misspecification.

In order to assess the power of different benchmarks to detect model misspecification, we perform a second set of Monte Carlo simulations in which we simulate data sets from a misspecified model but then estimate the model in Section 1. The misspecified model adds a cost of debt issuance to the model in Section 1, so if $p'-p>0$ and $p'>0$, then the firm must pay an issuance cost equal to $c\min(p'-p, p')$. To calculate a power curve for each of the tests we consider, we let c take eight evenly spaced values between 0.0025 and 0.04, performing a separate Monte Carlo simulation for each value. For each of these simulations, we calculate the power of three tests: the standard overidentification test, the external validity test, and the test from Nikolov and Whited (2014) that an element of $g(\boldsymbol{v}_{it},\theta)$ from the correctly specified model ($c=0$) equals the corresponding element from the misspecified model ($c>0$). For brevity, instead of reporting rejection rates for each element of $g(\boldsymbol{v}_{it},\theta)$, for these latter tests we report the maximum, median, and minimum test rejection rates.

Figure 4 presents the results for the EPF-based estimator for the large sample size of 75,000. Strikingly, the test of overidentifying restrictions has excellent power to detect even small amounts of misspecification. The power of the test shoots up to 1 even when the issuance cost is at the lowest level we consider, 0.0025. Although the external validity test does not exhibit such striking behavior, it nonetheless has good power to detect misspecification, with the rejection rates of 0.7 and 0.9 for issuance costs of 0.02 and 0.04. Finally, the *t*-tests for the individual elements of $g(\boldsymbol{v}_{it},\theta)$ also have good power, with the median and maximum rejection rates spiking to 1 even for small issuance costs up to 0.02. Even the minimum rejection rate eventually reaches 1 when the issuance cost is 0.04.

Figure 5 presents analogous results for the moments-based estimator for the large sample size of 75,000. Here, we observe somewhat lower power for the overidentification test and for the moment *t*-tests. Mirroring the results in Table 5, the external validity test over-rejects strongly even under the null, so the strong over-rejection away from the null is not surprising.

5. Robustness

Monte Carlo simulations have the obvious drawback that they pertain only to a specific parameter constellation for a specific model. To address this issue, we consider several extensions to our basic setup. In particular, we consider using additional moments in our moments-based estimation, an alternative transformation of the state variables in the EPF-based estimation, and a richer specification of the baseline polynomial approximation to the empirical policy function. Finally, we consider two alternate models: one adds fixed costs of capital adjustment to the basic constant-returns model in Section 1, and the other augments this baseline model by allowing for decreasing returns to scale.

First, we explore the result from Table 5 of the poor performance of the external validity test of the moments-based estimator. Our intuition for this result is that the moments-based estimation does not contain any of the covariances that define the empirical policy function estimates, so the parameters from the moments-based estimation struggle to replicate the empirical policy function coefficients. The question naturally arises whether adding covariance terms to the moments-based estimation can improve the performance of the external validity test. To address the question, we add three covariances to the moment list: $\mathrm{cov}(i,p)$, $\mathrm{cov}(i,d)$, and $\mathrm{cov}(p,d)$. This exercise also addresses the result from Tables 3 and 4 of the inferior performance of the moments-based estimator relative to the EPF-based estimator. Our intuition for this last result is that the lack of covariance moments in the moments-based estimation implies less identifying information to allow for precise parameter estimation.

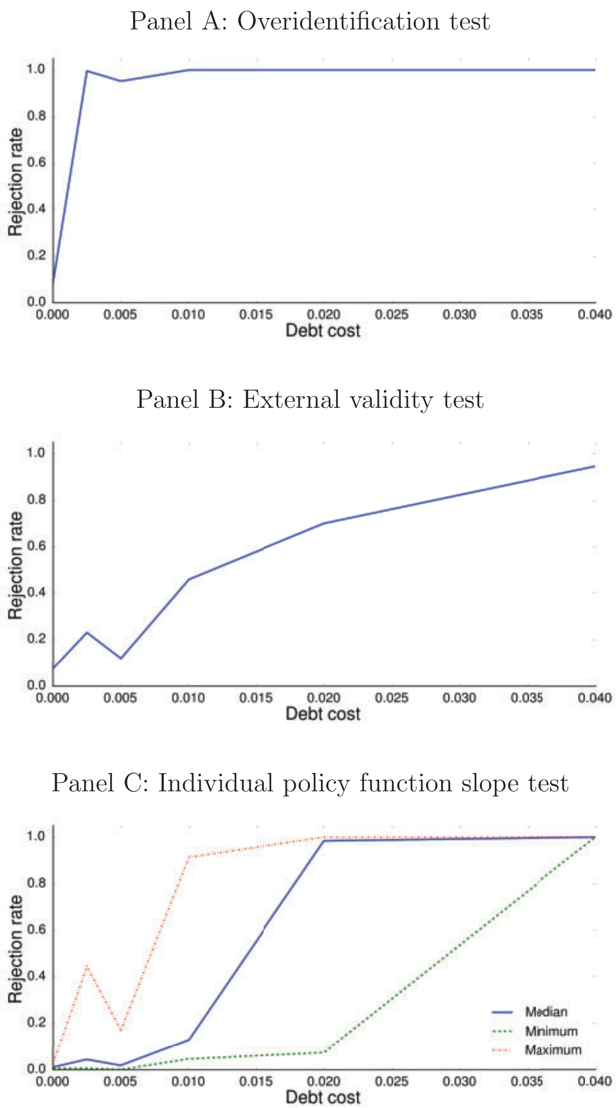

Figure 4
Specification test power in the EPF-based estimators
Indicated probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1, with a proportional cost of debt issuance, c. On the x-axis of each panel is the parameter c. On the y-axis is the rejection rate for the relevant test. All estimations use empirical policy functions as benchmarks. Panel A shows the actual size of a nominal 5% test of the model overidentifying restrictions, where the data are generated for a model in which $c>0$, but the model is estimated as if $c=0$. Panel B shows the actual size of a nominal 5% test of an external validity test whose null is that the moments (which are not used in the estimation) equal their true values. Panel C shows the power of the test from Nikolov and Whited (2014) that a policy-function slope from a model estimated with $c>0$ equals the same policy-function slope from a model in which c is restricted to zero. We report the maximum, median, and minimum rejection rates across our set of policy function slopes.

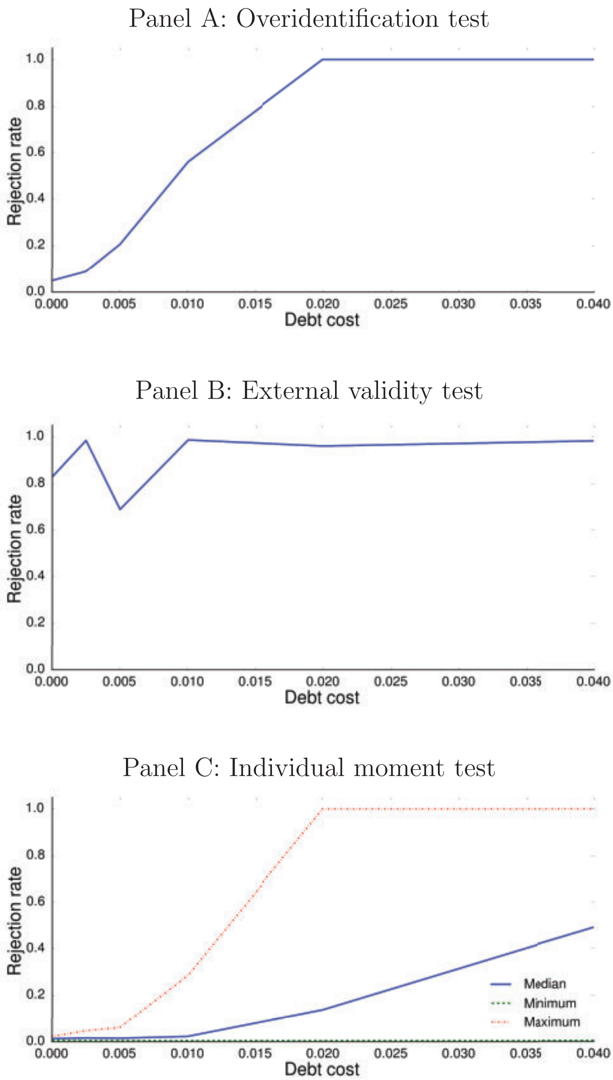

Figure 5
Specification test power in the moments-based estimators
Indicated probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1, with a proportional cost of debt issuance, c. On the x-axis of each panel is the parameter c. On the y-axis is the rejection rate for the relevant test. All estimations use moments as benchmarks. Panel A shows the actual size of a nominal 5% test of the model overidentifying restrictions, where the data are generated for a model in which $c > 0$, but the model is estimated as if $c = 0$. Panel B shows the actual size of a nominal 5% test of an external validity test whose null is that the moments (which are not used in the estimation) equal their true values. Panel C shows the power of the test from Nikolov and Whited (2014) that a moment from a model estimated with $c > 0$ equals the same moment from a model in which c is restricted to zero. We report the maximum, median, and minimum rejection rates across our set of moments.

Table 6
Monte Carlo comparison of simulation estimators: Covariance moments, high order EPFs, and transformed EPFs

Parameter	Extra covariance moments	High-order EPF-based	Transformed EPF-based
δ (depreciation rate)			
Average % bias	−0.002	0.001	0.000
RMSE %	0.029	0.004	0.007
Pr(t)	0.288	0.188	0.275
λ (equity issuance cost)			
Average % bias	0.068	−0.014	−0.200
RMSE %	1.210	0.429	0.857
Pr(t)	0.001	0.007	0.001
ξ (collateral parameter)			
Average % bias	−0.009	−0.012	0.012
RMSE %	0.240	0.044	0.101
Pr(t)	0.030	0.019	0.017
γ (investment adjustment cost)			
Average % bias	0.016	0.001	0.004
RMSE %	0.067	0.007	0.017
Pr(t)	0.101	0.055	0.077
Overidentification test rejection rate	0.071	0.090	0.097
External validity test rejection rate	0.710	0.081	0.086
Moment t-statistics:			
maximum rejection rate	0.028	0.015	0.030
median rejection rate	0.014	0.009	0.006
minimum rejection rate	0.011	0.004	0.005

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1. We consider three model extensions. For the first, we add extra covariance moments to the moments-based estimator. For the second, we consider a third-order polynomial approximation to the empirical policy function. For the third, we consider an alternative transformation of the state variables in the EPF-based estimator. All estimators use a clustered weight matrix. Bias and RMSE are expressed as fractions of the true coefficient. See Tables 3 and 5 for definitions.

In the first column of Table 6, we report the results of a Monte Carlo simulation of this augmented moments-based estimator. Here, we have gathered the parameter recovery and test size results into one table. All results are from estimators with a clustered weight matrix and a sample size of 75,000. Interestingly, for each parameter, the RMSE for this parameter lies between those for the moments-based and EPF-based estimators reported in Table 3. This result confirms our intuition concerning the identifying information in the extra moments. We also find that although the external validity test again over-rejects, the rejection rate is not as high as the 0.843 rate reported in Table 5. We conclude that adding these extra moments goes part of the way toward improving the performance of this test. However, because we are not adding all of the extra moments that define the empirical policy function estimates, this improvement is incomplete relative to the rejection rate of 0.079 for the EPF-based external validity test reported in Table 5.

Second, we consider using a more flexible approximation to the policy function. As discussed above, the issue at hand is whether using a closer-fitting approximation to the policy function improves finite sample performance.

On the one hand, a better fit ought to be able to better capture the dynamic interactions between the policy variables and thus be better able to identify the economic model parameters. On the other hand, closer-fitting approximations imply more auxiliary parameters to estimate and are thus likely to be higher variance, which in turn compromises identification of the economic model parameters.

Due to computational constraints, we can consider only a simple extension of our quadratic approximation of the empirical policy functions, specifically, the addition of third-order terms to the polynomial expansion in Equation (20). The results from this Monte Carlo simulation are in the second column of Table 6. The bias and RMSE figures for all parameters are slightly smaller than those in Table 3 from the EPF-based estimator with a clustered weight matrix and a second-order approximation to the policy function. Moreover, the sizes of both the t-tests and specification tests are closer to the nominal 5% value than the corresponding sizes in Tables 3 and 5. These results imply that for this simple extension, the improved accuracy of the policy function approximation improves finite sample performance.

Third, we explore the robustness of the EPF-based estimator to an alternative reparameterization of the policy and state variables. This issue is of interest for models in which the primitive state variables are unobservable and in which one must work with observable transformations of these unobservable variables. Although this issue is not relevant for our simple constant-returns model, we nonetheless consider a state variable transformation that has been widely used in the corporate finance literature. Specifically, instead of considering profitability and net debt (z and p), we express the empirical policy functions in terms of profitability and net worth: z and $\xi(1-\delta)+z-p$ (Cooley and Quadrini 2001; Hennessy and Whited 2007; Li, Whited, and Wu 2016). This transformation is of particular interest because model parameters appear in the expression for net wealth, just as a parameter appears in the transformation that we describe for a decreasing returns model below. The results from using this alternative transformation are in the third column of Table 6, where we find that the results are nearly identical to those from Table 3.

Next, we consider the possibility that the policy function might have sharp discontinuities. In particular, we augment our model from Section 1 to allow for the presence of fixed costs of capital adjustment that take the form of $\chi\mathcal{I}(i\neq 0)$. This model feature creates a sharp discontinuity in the policy function that relates investment, i, to the profit shock, z. The optimal investment policy is zero below a threshold for z, with a jump to a higher level of investment when z surpasses the threshold. Considering policy functions with discontinuities is of interest for two reasons. First, many models that have been studied in the literature contain similar nonconvexities (e.g., Cooper and Haltiwanger 2006). Second, and more importantly, the presence of a discontinuity might hinder parameter identification if the empirical approximation to the policy function does not capture the effects of the discontinuity.

Table 7
Monte Carlo comparison of simulation estimators: Model with fixed adjustment costs

Parameter	Second-order EPF-based	Third-order EPF-based
δ (depreciation rate)		
Average % bias	−0.003	−0.000
RMSE %	0.009	0.003
Pr(t)	0.025	0.057
λ (equity issuance cost)		
Average % bias	−0.144	0.063
RMSE %	1.111	0.566
Pr(t)	0.011	0.083
ξ (collateral parameter)		
Average % bias	−0.006	−0.000
RMSE %	0.008	0.005
Pr(t)	0.064	0.033
γ (investment adjustment cost)		
Average % bias	−0.061	−0.000
RMSE %	0.089	0.007
Pr(t)	0.030	0.030
θ (fixed adjustment cost)		
Average % bias	0.116	0.004
RMSE %	0.180	0.008
Pr(t)	0.032	0.040
Overidentification test rejection rate	0.049	0.107
External validity test rejection rate	0.016	0.043
Moment t-statistics:		
maximum rejection rate	0.011	0.017
median rejection rate	0.005	0.005
minimum rejection rate	0.001	0.000

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1, with one addition. We allow for the presence of fixed costs of capital adjustment that take the form of $\chi\mathcal{I}(i \neq 0)$. We consider two different benchmarks: a second-order approximation to the empirical policy function and a third-order approximation. Both estimators use a clustered weight matrix. Bias and RMSE are expressed as fractions of the true coefficient. See Tables 3 and 5 for definitions.

To explore this possibility, we start with our baseline model from Section 1 and calibrate the fixed-cost parameter χ at 0.025, which puts the jump in the policy function at the mean of z and sets the height of the jump at just over the depreciation rate. This setting implies that many large investment bursts associated with the policy function discontinuity occur in the simulated data, so the discontinuity is empirically relevant. We then perform a Monte Carlo simulation of two EPF-based estimators. Both use the optimal clustered weight matrix, but one uses a second-order approximation to the policy function, while the other uses a third-order approximation.

The results are in Table 7, where we report statistics for the four parameters in Table 3, as well as for the fixed-cost parameter, χ. For the second-order EPF-based estimator, we find near-zero bias and low RMSEs for all parameters, including χ. For the third-order EPF-based estimator, we find even smaller bias and RMSE figures for all parameters, including the bias and RMSE for χ.

The low bias and RMSEs for both of these EPF-based estimators at first appear counterintuitive, as neither a second- nor a third-order polynomial can typically capture a sharp discontinuity. However, nonconvexities in the optimization problem affect many other features of the policy function besides the obvious kink. For example, fixed investment adjustment costs affect the level of leverage and the response of investment to profitability. As long as these other features of the policy function are sufficient for parameter identification, then a tight-fitting policy function approximation need not be necessary. This finding reinforces the intuition in Gourieroux, Monfort, and Renault (1993) that the auxiliary model for indirect inference (the EPF approximation in our case) does not need to be a perfectly accurate description of the true distribution of the data. Instead, the auxiliary model simply needs to capture enough features of the data to identify the economic model parameters.

Finally, we consider an alternative model that augments the model from Section 1 by allowing for decreasing returns to scale. Specifically, we consider a production function that takes the form $zK^{\alpha}, \alpha < 1$. In this case, one cannot reduce the dimension of the state space by dividing all variables by K, so the state variable z is an unobservable shock. Therefore, to estimate this alternative model using the EPF-based estimator, we need to work with observable transformations of the state and control variables. For consistency with our earlier results, we use the natural transformation $\boldsymbol{w} = \{I/K, P'/K'\}$ and $\boldsymbol{x} = \{P/K, zK^{\alpha}/K\}$. Finally, we simplify the process governing z by setting the mean of the log process, μ, to zero. This specification is standard in the literature (e.g., Gomes 2001; Hennessy and Whited 2005).

Using this model in our Monte Carlo simulations is interesting for two reasons. First, the dimension of the state space is larger, and it is worth investigating whether the dimension of the state space matters for estimator performance. Second, models based on decreasing returns technology are much more widely used in the literature than models based on constant returns technology. Indeed, they are a workhorse in the macroeconomics and corporate finance literatures (e.g., Gomes 2001; Hennessy and Whited 2005), so a Monte Carlo based on this class of models is of broad interest.

In order to calibrate the Monte Carlo simulations, we need a baseline model parameterization, which we obtain by estimating the model using our second-order EPF-based estimator. For brevity, we do not report the full estimation results, but Table 8 lists the parameters, all of which are significantly different from zero in our estimation. Table 8 also contains the results from two Monte Carlo simulations, where we consider a moments-based estimator and a second-order EPF-based estimator, both of which use a clustered optimal weight matrix. For these Monte Carlos, we only estimate four of the model parameters, treating α, σ and ρ as given. In addition, computational constraints limit us to considering only 100 trials.

Even when we use the more elaborate decreasing returns model, Table 8 shows that the performance of both estimators is nearly identical to the performance documented in Table 3 for the constant returns model. In particular, the large RMSE for the moments-based estimator of the issuance

Table 8
Monte Carlo comparison of simulation estimators: Decreasing returns to scale model

Parameter	Moments-based	EPF-based
δ (depreciation rate)		
Average % bias	−0.048	0.004
RMSE %	0.076	0.025
Pr(t)	0.010	0.060
λ (equity issuance cost)		
Average % bias	−2.448	−0.745
RMSE %	3.275	1.492
Pr(t)	0.000	0.020
ξ (collateral parameter)		
Average % bias	−0.012	−0.005
RMSE %	0.036	0.019
Pr(t)	0.120	0.050
γ (adjustment cost)		
Average % bias	1.293	−0.081
RMSE %	1.768	0.594
Pr(t)	0.010	0.040
Overidentification test rejection rate	0.210	0.080
External validity test rejection rate	0.910	0.170
Moment t-statistics:		
maximum rejection rate	0.310	0.040
median rejection rate	0.065	0.015
minimum rejection rate	0.010	0.000

Indicated expectations and probabilities are estimates based on 1,000 Monte Carlo samples of size 75,000. The samples are generated from the model in Section 1, with two modifications. We allow for decreasing returns to scale, with the profit function curvature given by α. We also omit investment adjustment costs. We consider two different benchmarks: traditional moments and empirical policy functions. Both estimators use a clustered weight matrix. Bias and RMSE are expressed as fractions of the true coefficient. See the captions to Tables 3 and 5 for definitions.
True estimated parameter values: $\delta = 0.144$, $\xi = 0.521$, $\gamma = 2.173$, and $\lambda = 0.014$.
True fixed parameter values: $\alpha = 0.528$, $\rho = 0.297$, and $\sigma = 0.489$.

cost parameter, λ, and the external validity test over-rejection for the moments estimator are qualitatively similar across the two tables. This similarity arises because mapping from moments or policy function coefficients to the underlying structural parameters is similar in these two classes of models. For example, average leverage (or the intercept in the debt policy function) maps strongly into the collateral parameter, ξ, in both models.

6. Conclusion

This paper contributes in three ways to the methods that help us understand the interface between models and data. First, we assess the finite sample properties of popular simulation estimators in a setting relevant to researchers who use micro data to evaluate dynamic models. The results are mostly positive. The estimators produce unbiased coefficient estimates, and when we use an optimal weight matrix for the estimation of the model parameters, the specification tests associated with the estimator and the t-tests associated with the parameter estimates are close to correctly sized. Most importantly, the specification tests associated with the simulation estimators we consider have excellent power to detect misspecification. Given the paucity of work on the finite sample properties of simulation estimators, these results provide important guidance to applied researchers interested in using these estimators.

Second, we introduce a specification test that holds dynamic models to a high standard by assessing their ability to reconcile features of the data not used in their estimation. We find that this test can over-reject for a simple moments-based estimator but that it is close to correctly sized in the case of our EPF-based estimator. In this case, it also has excellent power.

Third, we introduce a set of statistical benchmarks to use for the quantitative evaluation of dynamic models. The benchmarks are empirical policy functions—that is, the empirical counterparts of the policy functions from these models. We argue that these benchmarks are intuitive, robust, and theoretically motivated. We then demonstrate how to use these benchmarks as a basis for the estimation of model parameters using indirect inference.

The biggest advantage of using empirical policy functions as a basis for the estimation of dynamic models is discipline. In particular, the use of policy functions alleviates the common concern with using data moments that moments can be cherry-picked to steer the estimation results a certain way. Another advantage is that policy functions characterize the solutions to all dynamic models, so using them as benchmarks for evaluation makes sense from an economic perspective. The combination of these two advantages implies that estimators based on empirical policy functions can be used to compare models, as long as the models describe the same policies.

Model comparison has been a neglected feature of the evaluation of dynamic models. Although many studies (e.g., Whited, 1992; Kadiyali, Sudhir, and Rao 2001; Nikolov and Whited 2014) compare nested models, few compare nonnested models that describe similar policy variables. We therefore view using empirical policy function benchmarks to compare both nested and nonnested models as a natural extension of the research we have presented here.

Appendix

This Appendix sketches the standard test statistics that we use. The covariance matrix for the parameter vector, θ, is given by:

$$\frac{1}{nT}\left(1+\frac{1}{S}\right)(G'\hat{W}G)^{-1}G'\hat{W}\hat{\Omega}\hat{W}G(G'\hat{W}G)^{-1}, \tag{A1}$$

in which $G \equiv \partial g(\boldsymbol{v}_{it},\theta)/\partial\theta$. This expression is a standard generalized method of moments (GMM) parameter variance formula with a correction for simulation error, which is given by $\left(1+\frac{1}{S}\right)$. As is standard, when $\hat{W}=\hat{\Omega}^{-1}$, Equation (A1) reduces to:

$$\frac{1}{nT}\left(1+\frac{1}{S}\right)(G'\hat{\Omega}^{-1}G)^{-1}. \tag{A2}$$

Second, the variance of the vector $g(\boldsymbol{v}_{it},\theta)$ is given by:

$$\text{var}(g(\boldsymbol{v}_{it},\theta))=\frac{1}{nT}\left(1+\frac{1}{S}\right)(I-G(G'WG)^{-1}G'W)\hat{\Omega}(I-G(G'WG)^{-1}G'W). \tag{A3}$$

Third, the test of overidentifying restrictions when one uses an arbitrary weight matrix, W, is given by:

$$\frac{nTS}{1+S} g(\boldsymbol{v}_{it},\theta)' \mathrm{var}(g(\boldsymbol{v}_{it},\theta))^{+} g(\boldsymbol{v}_{it},\theta), \tag{A4}$$

in which + indicates a pseudo-inverse. In the case in which W is the optimal weight matrix, this test takes the familiar form

$$\frac{nTS}{1+S} g(\boldsymbol{v}_{it},\theta)' \hat{\Omega}^{-1} g(\boldsymbol{v}_{it},\theta) \tag{A5}$$

The last two tests have degrees of freedom equal to the dimension of $g(\boldsymbol{v}_{it},\theta)$ minus the dimension of θ.

References

Abel, A. B., and J. C. Eberly. 1994. A unified model of investment under uncertainty. *American Economic Review* 84:1369–84.

Arellano, M., and S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58:277–97.

Bajari, P., C. L. Benkard, and J. Levin. 2007. Estimating dynamic models of imperfect competition. *Econometrica* 75:1331–70.

Cochrane, J. H. 2005. *Asset pricing,* rev. ed. Princeton, NJ: Princeton University Press.

Cooley, T. F., and V. Quadrini. 2001. Financial markets and firm dynamics. *American Economic Review* 91:1286–310.

Cooper, R. W., and J. C. Haltiwanger. 2006. On the nature of capital adjustment costs. *Review of Economic Studies* 73:611–33.

DeAngelo, H., L. DeAngelo, and T. M. Whited. 2011. Capital structure dynamics and transitory debt. *Journal of Financial Economics* 99:235–61.

Eisenhauer, P., J. J. Heckman, and S. Mosso. 2015. Estimation of dynamic discrete choice models by maximum likelihood and the simulated method of moments. *International Economic Review* 56:331–57.

Erickson, T., and T. M. Whited. 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. *Econometric Theory* 18:776–99.

———. 2012. Treating measurement error in Tobin's *q*. *Review of Financial Studies* 25:1286–329.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen. 1988. Financing constraints and corporate investment. *Brookings Papers on Economic Activity* 1:141–206.

Gala, V., and J. F. Gomes. 2016. Investment without q. Manuscript, University of Pennsylvania.

Gallant, A. R., and G. Tauchen. 1996. Which moments to match? *Econometric Theory* 12:657–81.

Gomes, J. F. 2001. Financing investment. *American Economic Review* 91:1263–85.

Gormley, T. A., and D. A. Matsa. 2014. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies* 27:617–61.

Gourieroux, C. S., A. Monfort, and E. Renault. 1993. Indirect inference. *Journal of Applied Econometrics* 8:S85–S118.

Hansen, L. P., J. Heaton, and A. Yaron. 1996. Finite-sample properties of some alternative GMM estimators. *Journal of Business and Economic Statistics* 14:262–80.

Hennessy, C. A., and T. M. Whited. 2005. Debt dynamics. *Journal of Finance* 60:1129–65.

———. 2007. How costly is external financing? Evidence from a structural estimation. *Journal of Finance* 62:1705–45.

Ingram, B. F., and B.-S. Lee. 1991. Simulation and estimation of time series models. *Journal of Econometrics* 47:197–205.

Kadiyali, V., K. Sudhir, and V. R. Rao. 2001. Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *International Journal of Research in Marketing* 18:161–86.

Li, S., T. M. Whited, and Y. Wu. 2016. Collateral, taxes, and leverage. *Review of Financial Studies* 29:1453–500.

Michaelides, A., and S. Ng. 2000. Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators. *Journal of Econometrics* 96:231–66.

Nikolov, B., and T. M. Whited. 2014. Agency conflicts and cash: Estimates from a dynamic model. *Journal of Finance* 69:883–1921.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Rampini, A. A., and S. Viswanathan. 2013. Collateral and capital structure. *Journal of Financial Economics* 109:466–92.

Schroth, E., G. Suarez, and L. A. Taylor. 2014. Dynamic debt runs: Evidence from a structural estimation. *Journal of Financial Economics* 112:164–89.

Shanken, J., and G. Zhou. 2007. Estimating and testing beta pricing models: Alternative methods and their performance in simulations. *Journal of Financial Economics* 84:40–86.

Smith, Jr., A. A. 1993. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics* 8:63–84.

Stokey, N. L., R. E. Lucas, and E. C. Prescott. 1989. *Recursive methods in economic dynamics*. Cambridge, MA: Harvard University Press.

Strebulaev, I. A., and T. M. Whited. 2012. Dynamic models and structural estimation in corporate finance. *Foundations and Trends in Finance* 6:1–163.

Taylor, L. A. 2010. Why are CEOs rarely fired? Evidence from structural estimation. *Journal of Finance* 65:2051–87.

van Binsbergen, J., and C. Opp. 2016. Real anomalies: Are financial markets a sideshow? Manuscript, University of Pennsylvania.

Warusawitharna, M., and T. M. Whited. 2016. Equity market misvaluation, investment and finance. *Review of Financial Studies* 29:603–54.

Whited, T. M. 1992. Debt, liquidity constraints, and corporate investment: Evidence from panel data. *Journal of Finance* 47:1425–60.