BOSTON UNIVERSITY

Jean-Jacques Forneron
Assistant Professor
Department of Economics,
Boston University

270 Bay State Road
Boston MA, 02215
E-mail: jjmf@bu.edu
jjforneron.com

Research Statement

My research focuses on theoretical econometrics, with contributions to methodology and computation, particularly for estimation and inference of structural economic models. Broadly, my research spans three main areas: (1) robustness to global misspecification in structural models, (2) detecting and dealing with identification failure, and (3) the study of numerical optimization from an econometric perspective.

1. Robust Methods for Structural Estimation and Inference

Structural models are routinely used to test economic theories and conduct counterfactual analyses with non-experimental data. The model is often the result of seeking to balance an accurate description of real-world behaviour with some amount of mathematical and/or numerical tractability. In a series of papers, I propose methods to increase the robustness of structural estimates to the type of misspecification that can arise when the researcher estimates a more interpretable and tightly parameterized structural model.

In "A Sieve-SMM Estimator for Dynamic Models" (Forneron, 2023c), I consider a setting where the model is correctly specified up to distributional assumptions. The motivation is that although there is evidence of skewness in some macroeconomic shocks, the standard choice for estimation is to assume Gaussian shocks. Under these parametric assumptions, we can evaluate the likelihood to perform either frequentist or Bayesian inference. In this work, to relax this Gaussian distributional assumption, I propose and study a simulation-based semi-nonparametric estimator for the low-dimensional structural parameters and the distribution of the structural shocks in non-linear state-space models. Relaxing strong distributional assumptions can be particularly useful in resolving some well-known 'puzzles' in macroeconomics and finance. I illustrate this by estimating a bond pricing model in a production economy; a setting where it is particularly challenging to simultaneously fit macroeconomic and financial data with a fully parametric specification. I find that the skewness and tail risks in productivity and inflation implied by the flexible distribution estimates ease the tensions in fitting the data and produces more credible parameter estimates.

"Fitting Dynamically Misspecified Models: An Optimal Transportation Approach" (Forneron and Qu, 2024) continues this work by considering structural models where the model's dynamics cannot match those of the data. We show that, because standard filtering algorithms do not enforce model constraints, filtered variables may not have a clear interpretation within the model framework. For example, shocks that should be independent can be correlated. This can make the interpretation of parameter estimates ambiguous. We propose and study a recursive optimal transportation (OT) algorithm which constructs a model-consistent sample that is closest to the original data in mean-squared error. Within this model-consistent dataset, filtered variables and parameter estimates have a model-consistent interpretation. To evaluate the fit of the model, the discrepancy between the data and the model-consistent sample can be used to formally test the model specification on all variables or a particular subset of interest. For linear processes, the algorithm has closed-form and only involves standard matrix operations - a significant computational improvement relative to related applications of optimal transportation in statistics and machine learning. We derive large sample properties for the estimates and specification test. We apply our methodology to the Lubik-Schorfheide and Smets-Wouters macroeconomic DSGE models as well as an affine term-structure model for bond yields. For the Smets-Wouters model, the fit for consumption is rejected, as the fitted variance and persistence are too low, and the model cannot match the size of recessions. In a validation exercise with the same model, we also find that our Optimal Transport Filter more accurately recovers variables that were available but intentionally excluded from estimation.

In an ongoing work, my colleague Zhongjun Qu and I aim to extend our methodology to find numerically efficient algorithms for more general non-linear and non-Gaussian state-space models. This would enable users to evaluate the fit of a broader class of nonlinear asset pricing and macroeconomic models on historical data; including non-linear DSGE and stochastic volatility models.

"Occasionally Misspecified" (Forneron, 2023b) takes a different perspective on the same problem. I consider the case where the model provides a good fit for most but not all observations. This can occur when the data includes some periods of temporary policy experimentation or extreme events. These can be thought of as outliers which could invalidate the parameter estimates' external validity, if these events were to never re-occur. A natural solution is to use a robust estimator as it limits the influence of these outliers. This type of regularization produces estimates that are generally inconsistent when the underlying data is asymmetric or asymmetrically contaminated. Many variables studied in empirical work are asymmetric, or in other words skewed: income, quantity, price, size of cities, to name a few. Yet, the issue of asymmetry bias in robust estimation has received limited attention. Building on ideas from non-parametric estimation, I propose a framework where a tuning parameter controls the tradeoff between contamination and regularization bias for a simple estimator. I derive its finite and large sample properties, then I propose a bias correction technique which permits robustness to higher proportions of contamination without introducing asymptotic regularization bias. An oracle equivalence result shows that the estimates are asymptotically equivalent to an oracle who first removes all outliers from the sample and then computes estimates using only the 'good' observations. In the main application, I revisit a well-known price puzzle which finds a price increase in reaction to an increase in interest rate, when using a recursive VAR to identify monetary policy shocks. I find that observations associated with the Volcker era, a sudden but temporary shift in the conduct of monetary policy, are highly influential. Robust estimates reduce their influence but inherit some bias given the skewness of the data. The bias-corrected estimates reduce the influence of the same datapoints and recover theory-consistent predictions.

2. Detecting Identification Failure and Identification-Robust Inference

Standard asymptotic inferences on structural parameters using a t or Wald test require said parameters to be both uniquely and locally identified. If these conditions fail or nearly fail, identification-robust inference should be used. In empirical work, identification is often assumed as it is challenging to verify analytically for nonlinear models.

In "Detecting Identification Failure in Moment Condition Models" (Forneron, 2024a), I propose a numerical approach to determine whether parameters are both globally and locally identified, or instead are set-identified or weakly identified. To this end, I introduce a quasi-Jacobian matrix, computed as the solution of a linear program, which is singular when at least one of the local and global identification conditions fails. When both conditions hold, it recovers the usual Jacobian matrix. Compared to earlier proposals, the matrix is informative about the parameters involved in the lack of identification. Leveraging this information, I propose a two-step procedure to perform identification robust subvector inference. The test is more powerful than simple projection inference when some or all nuisance parameters are strongly or semi-strongly identified. Using my method, I find that identification fails, or nearly fails, for several parameters in the long-run risks model. Although the test is more conservative than standard methods, I also find that inferences on the agent's utility function can be informative about their attitude towards risk and their time preference.

Leveraging techniques from the paper above and the Sieve-SMM estimator, an ongoing work with my colleague Zhongjun Qu builds a general method to determine whether a model is parametrically identified, locally and globally. Our focus is on intractable models, where the likelihood does not have closed-form and is computationally intensive to evaluate; e.g. using the particle filter. To this end, we leverage a distance that is (1) easy and numerically cheap to compute via simulation, and such that (2) if a model is locally and globally identified under this metric, then the same holds under the Kullback–Leibler divergence used for Likelihood estimations and Bayesian inference. The methodology is applicable to a wide range of non-linear intractable models that arise in e.g. macroeconomics, finance, and labour economics, where the likelihood or the moments used for estimation are computed using Monte Carlo simulation.

3. Numerical Optimization: an Econometric Perspective

Numerical optimization is an essential aspect of empirical work for economists working with non-linear structural models in many fields, including industrial organization, trade, labour, and macroeconomics. While much of the econometric literature is devoted to proposing estimators and studying their statistical properties, I find that much less attention has been devoted to the methods that are used to compute said estimators in practice.

Optimization-based sampling. In a series of papers, starting with my PhD thesis, I propose and study sampling algorithms that builds on optimization techniques to perform both estimation and inference at a reduced computational time.

In "A Likelihood-Free Reverse Sampler of the Posterior Distribution" (Forneron and Ng, 2016), we propose a sampling algorithm for Approximate Bayesian Computation (ABC) which, in contrast to existing methods, does not suffer from the curse of dimensionality. ABC is a likelihood-free Bayesian framework that is particularly popular in the sciences, where a posterior is computed from a summary statistic rather than the full sample. This is similar to methods used in econometrics: the simulated method of moments (SMM) and indirect inference (II). The curse of dimensionality, in the number of moments, implies that approximating an infeasible posterior with higher accuracy comes with increased computational burden. The sampler we propose is optimization-based which avoids the curse of dimensionality, the computation burden is much lower, and posterior inferences are more accurate. We illustrate the relative performance with a simple ARMA estimation and a model with precautionary savings. In "The ABC of Simulation Estimation with Auxiliary Statistics" (Forneron and Ng, 2018) we compare the theoretical properties of this ABC algorithm with frequentist econometric methods: SMM and II, as well as quasi-Bayesian estimators that can be constructed from SMM and II. We illustrate the differences in computation and statistical properties for the estimation of a dynamic panel model with fixed effects.

The idea of using optimization for sampling is further explored in "Estimation and Inference by Stochastic Optimization" (Forneron, 2024b). Building on ideas from the stochastic gradient descent literature, this paper proposes algorithms with fixed tuning parameters in which the loss function is resampled at each iteration. For convex problems, the random iterates can be used as draws where their average is an asymptotically valid estimator, and their distribution provides valid bootstrap inference, after rescaling for subsample size and scaling parameters. Similar to Bayesian sampling algorithms, this allows researchers to perform point estimation and bootstrap inference simultaneously. I illustrate the results and numerical performance using simulation for a dynamic discrete choice model. A large-scale probit estimation of a gravity model using cross-country importer-exporter data illustrates good performance relative to gradient-based MCMC algorithms used in Bayesian inference. Finally, in a replication I show that the method finds accurate estimates and confidence intervals for the trade cost of salt in India between 1861-1930 in 6 hours on a desktop computer, compared to several days on a 100-core cluster environment used the original study. In "Estimation and Inference by Stochastic Optimization: Three Examples" (Forneron and Ng, 2021), we illustrate the methodology on three examples and show statistical efficiency gains for simulation-based estimators (SMM, II), similar to our own ABC algorithm but at a lower computational cost.

Non-Convex Estimation of Moments Condition Models. The Generalized Method of Moments (GMM) and Simulated Method of Moments (SMM) estimators are particularly popular with empiricists working on structural economic models in industrial organization, labor, macroeconomics, and trade. In practice, practitioners often report that estimation is fairly challenging because of non-convexity and/or non-smoothness in the GMM or SMM loss function. This particular issue has been the focus of some of my more recent work.

In "Noisy, Non-Smooth, Non-Convex Estimation of Moment Condition Models" (Forneron, 2023a), I propose an optimization algorithm for GMM and SMM estimations that only satisfy textbook econometric assumptions and study its finite and large sample properties. The algorithm runs concurrently a global and local search step. I show that, after finitely many iterations, the algorithm converges exponentially fast to an estimator with standard large-sample properties (consistent and asymptotically Gaussian). Compared to popular alternatives in econometrics and machine learning, the non-smooth loss function is not smoothed. Instead, the local step relies on a combination of a smoothed Jacobian and non-smooth moments, which leads to favorable statistical properties: the estimates are asymptotically unbiased under much weaker assumptions than required by other methods. This allows the user to set the tuning parameters to improve numerical performance without degrading the first-order asymptotic properties of the estimates. The algorithm I propose is shown to be globally convergent for non-convex problems. This is in sharp contrast with many proposals that study convex smoothed loss functions. Simulations for a dynamic discrete choice model and a discretized model with precautionary savings show that the algorithm outperforms several packaged global optimizers, at a significantly lower computational cost. An empirical application to joint retirement decisions for couples further illustrates the performance and accuracy of the algorithm I propose. While the theoretical and empirical results indicate good performance, a curse of dimensionality remains which limits the scale of estimations.

We consider this last issue more specifically in "Convexity Not Required: Estimation of Smooth Moment Condition Models" (Forneron and Zhong, 2025). Besides standard assumptions, we introduce an additional condition that is sufficient for gradient-based optimizers to be globally convergent. We show that our condition is weaker than strong convexity and star convexity of the GMM loss function, as well as strong monotonicity or injectivity conditions for just-identified moments. Our results provide some useful guidance for practitioners, for instance: if the moments are just-identified and strongly injective, then Gauss-Newton is globally convergent. This result is invariant to strongly injective reparameterizations and is thus fairly robust. There is a close connection with the Polyak-Łojasiewicz condition, a popular relaxation of convexity, if the moments can be set exactly equal to zero. For over-identified models, this is generally not the case in finite samples. Nevertheless, the results are robust to moderate misspecification, a case where the moments cannot be set exactly equal to zero asymptotically. Importantly for practitioners, our results indicate that when misspecification becomes arbitrarily large, local optima can arise in the GMM loss function. This negates global convergence results and implies a much more complex minimization problem. We revisit the estimation of a random coefficient demand model for cereal, which alerted researchers to a number of numerical issues associated with the estimation of highly nonlinear models. As in the original study, we find mixed performance for off-the-shelf methods, including convex optimizers. In contrast, our preferred Gauss-Newton algorithm converges systematically from a wide range of starting values. The estimation of a DSGE model with endogenous R&D confirms these findings in a macroeconomic estimation setting.

References

Forneron, J.-J. (2023a): "Noisy, non-smooth, non-convex estimation of moment condition models," *Conditionally Accepted at Quantitative Economics.*

——— (2023b): "Occasionally misspecified," *arXiv preprint arXiv:2312.05342.*

——— (2023c): "A sieve-SMM estimator for dynamic models," *Econometrica*, 91, 943–977.

——— (2024a): "Detecting identification failure in moment condition models," *Journal of Econometrics*, 238, 105552.

——— (2024b): "Estimation and inference by stochastic optimization," *Journal of Econometrics*, 238, 105638.

Forneron, J.-J. and S. Ng (2016): "A likelihood-free reverse sampler of the posterior distribution," in *Essays in Honor of Aman Ullah*, Emerald Group Publishing Limited, 389–415.

——— (2018): "The ABC of simulation estimation with auxiliary statistics," *Journal of Econometrics*, 205, 112–139.

——— (2021): "Estimation and Inference by Stochastic Optimization: Three Examples," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 111, 626–630.

Forneron, J.-J. and Z. Qu (2024): "Fitting Dynamically Misspecified Models: An Optimal Transportation Approach," *arXiv preprint arXiv:2412.20204.*

Forneron, J.-J. and L. Zhong (2025): "Convexity not required: Estimation of smooth moment condition models," *arXiv preprint arXiv:2304.14386.*