November 14, 2010

Professor Whitney Newey
Co-Editor, Econometrica
Department of Economics
MIT, USA

Dear Co-Editor Newey,

I am writing to resubmit the paper "Constrained Optimization Approaches to Estimation of Structural Models" (co-authored with Kenneth L. Judd), Manuscript 7925, for the second-round review at Econometrica. I am grateful to the referees for their excellent reports. Based on the Co-Editor's and the referees' suggestions, we have substantially revised and reorganized our paper. I hope that in the revised version we have carefully addressed all comments from the Co-Editor's decision letter and the referees' reports. In this letter I briefly explain the details of our revision and give the response to the remarks in the Co-Editor's letter. Responses to the referees' comments are given in the attached note.

The major changes in the revised version of the paper are:

(*i*). We add a new section (Section 2) on overview of numerical optimization methods.

(*ii*). We prove the equivalence between NFXP and MPEC for the Zucher model in Proposition 2. We add a new Monte Carlo study on the Rust's bus example to compare the performance of NFXP and MPEC.

(*iii*). We replaced the example on Bertrand pricing game in the previous version by the discrete-choice games of incomplete information suggested by Referee 4. We also prove the equivalence of NFXP and MPEC for this example in Proposition 3. We conduct a Monte Carlo study on this example to compare the performance of two-step estimators, NPL and MPEC.

(*iv*). We deleted Section 2, Section 4 and Section 5.4 in the previous version.

Since we have substantially revised our paper, we summarize the major changes in the previous and current version in Table 1 and 2, respectively.

Table 1: Changes on the Previous Version.

| Section | Comment |
|---|---|
| 1. Introduction | Revised |
| 2. Current Optimization Methods in Econometrics | Deleted. |
| 3. MPEC Approach to Estimation | Revised. |
| 4. MPEC Applied to a Simple Demand Example | Deleted. |
| 5. MPEC applied to Zurcher | (*i*) Section 5.1 to 5.3 are revised as part of Section 4 in the new version.<br>(*ii*) Section 5.4 is deleted. |
| 6. The MPEC Approach to Games | (*i*) We remove the he Bertrand pricing game example and replace it by the Prisoner Dilemma example suggested by Referee 4.<br>(*ii*) It is now Section 5 in the new version. |
| 7. The MPEC Approach and the Method of Moments | Deleted. |
| 8. Conclusion | Revised |

Table 2: Contents in the Current Version.

| Section | Comment |
|---|---|
| 1. Introduction | |
| 2. Overview of Numerical Optimization Methods | New material. |
| 3. MPEC Approach to Estimation | Include a result on the asymptotic distribution of the ML estimates with MPEC.. |
| 4. Single-Agent Dynamic Discrete Choice Models | (*i*) Based on Section 5.1 to 5.3 of the previous version.<br>(*ii*) Proof on the equivalence of NFXP and MPEC.<br>(*iii*) A new Monte Carlo study to compare the performance of NFXP and MPEC. |
| 5. Estimation of Games of Incomplete Information | (*i*) An example of discrete-choice game of incomplete information suggested by Referee 4.<br>(*ii*) Proof on the equivalence of NFXP and MPEC.<br>(*iii*) A numerical example that shows the discontinuity in the likelihood function of NFXP.<br>(*iv*) A new Monte Carlo study to compare the finite sample performance of two-step estimators, NPL and MPEC. |
| 6. Conclusion | |
| Appendix B. Automatic Differentiation | Exacted from Section 2 of the previous version. |

Below I repeat parts of the remarks from the Co-Editor's letter and then describe the resulting changes in the revised version of the paper.

**Editor's comments**:

**1.** *I endorse the suggestion to run a head-to-head comparison with a modern Nested Fixed Point (NFXP) implementation. Also, please provide at least one additional comparison with existing approaches, in the context of estimation of a game. Referee 4 provides one suggestion.*

We conduct a Monte Carlo study to compare the computational performance of NFXP and MPEC in Section 3 of the revised version. We follow the specifications used to report the estimates in Table X in Rust (1987). We use the reported parameter values in Table X to generate synthetic datasets. In our experiment, we also vary the discount factor $\beta$ and investigate the performance of NFXP and MPEC under various discount factor $\beta$. We provide two implementations of the MPEC estimator. The first implementation is conducted in AMPL, a modeling language which uses automatic differentiation to compute the exact first-order and second-order derivatives and analyzes the sparsity patterns of constraint Jacobian and Hessian. The second implementation of MPEC is conducted in Matlab, which we provide the hand-coded first-order derivatives and sparsity pattern of constraint Jacobian. The implementation of NFXP is conducted in Matlab with hand-coded first-order derivatives. We believe that our implementations of NFXP and MPEC in Matlab provide a head-to-head and fair comparison of the two algorithms.

The results show that MPEC is faster than NFXP. As expected, the speed of NFXP depends on the discount factor $\beta$ because NFXP uses contraction mapping iteration to solve the inner-loop problem. The computational time and the number of contraction mapping iteration needed in NFXP increase when the discount factor $\beta$ increases.

We also provide a Monte Carlo study on estimation of a discrete-choice game of incomplete information suggested by Referee 4. We follow the example studied in the lecture notes provided by Referee 4 and conduct three experiments, in which we use different type of equilibria in the data generating process. We show in an example of one market (Example 3 in Section 5.2.2 in p. 29) that the likelihood function of NFXP in the parameter space is discontinuous (in Figure 3, p. 30). In our Monte Carlo study, we consider a model of multiple markets, where a market is defined by the players' observed types. In the data generating process, we follow the Editor's suggestion to focus on the case that for each market, only one equilibrium is played in the data. However, we allow for different equilibria being played in different markets (with different observed types).

We compare the performance of two-step pseudo maximum likelihood (2S-PML), two-step least squares (2S-LS), nested pseudo likelihood (NPL) and MPEC. We conduct three experiments. In Experiment 1 and 2, we use best-reply stable equilibria to generate data; in Experiment 3, we allow for best-reply unstable equilibria in the data generating process.

The results (reported in Section 5.3 of the revised version) show that MPEC performs well across all three experiments. Two-step estimators perform well in Experiment 1 and 2, but are severely biased in Experiment 3. NPL performs well in Experiment 1, but is biased and has large standard error in one parameter in Experiment 2. NPL fails to converge after 500 NPL iterations in all 100 replications in Experiment 3. In an another Monte Carlo experiment, which we do not include in the revised version of the paper, NPL converges to wrong estimates.

**2.** *.... Also, it would be good to know what problems might come up with MPEC. For instance, could having enough computer memory be a problem with all those constraints?*

In Section 2, we describe a few features that are incorporated in the implementation of modern constrained optimization algorithms. One feature is to explore the sparsity pattern in constraint Jacobian and Hessian. In general, the computational time and memory required in constrained optimization algorithms is on the order of nonzero elements in constrained Jacobian and Hessian. It is true that for models with dense matrices, the memory requirement for MPEC would be an issue. However, in many applied models such as Rust's bus example or dynamic games, the Jacobian and Hessian are highly sparse. As we discuss in Section 4.2 (p. 16 and p. 17), the number of nonzero elements in constraint Jacobian and Hessian in Rust's Zucher model is in the order of $dim(\theta)+dim(EV)$, where $dim(\theta)$ is the number of structural parameters and $dim(EV)$ is the number of grid points on the mileage state space. Taking advantage of the sparsity structure in these models makes implementing MPEC computationally feasible even when the number of variables and constraints are large.

**3.** *One of the referees suggested that more information about how to actually use the software being provided. I think that would be great for the supplementary material website.*

We certainly will make our code available to the public and for the supplementary material website at Econometrica. We are in the process of creating a webpage that contains the code that we use in our Monte Carlo experiments. In the next few days, the code will be available at

`http://faculty.chicagobooth.edu/che-lin.su/research/code.html`

**4.** *Another important issue is multiple equilibria. ... Given this concern, and given that the primary focus of this paper is on computation, it seems best to drop the attempt to do something more about multiple equilibria, and just focus on computational methods for models from the established literature, where the assumption that the data corresponds to one equilibrium is maintained.*

We followed the Co-Editor's suggestion on this. In Assumption 2 (in Section 5.2.1, p 28) we state that "For each market, only one equilibrium is played in the data. However, equilibria played across different markets are different." We maintain this assumption when generating data for our Monte Carlo study in Section 5.3.

**5.** *It would also be good to discuss efficiency issues somewhat and to be more precise about the bootstrap.*

We include a result on the asymptotic distribution of the maximum likelihood estimates based the MPEC approach in Proposition 1 on p. 10 in Section 3.2. We do not prove this result because it follows directly from Theorem 2 in Aitchison and Silvey (1958). In Section 4.5, we explain that the sampling procedure used in our Monte Carlo study in Section 5.3 is a parametric bootstrap procedure to generate simulated samples and to obtain standard errors.

**6.** *One question raised by the referees is whether MPEC is simply a different way to compute the MLE or whether it changes the estimator. It would be good to clarify this, e.g., in the context of the Zucher model,where a referee asks if the estimators are different.*

MPEC is simply a different way to compute the MLE. It does not change the estimator. We state this result formally in Proposition 2. in Section 4.3 (p. 17) and provide the proof in Appendix A. The same statement about the equivalence of MPEC and NFXP is made in Proposition 3 (p. 32) in Section 5.2.3 for the discrete-choice game of incomplete information.

**7.** *The referees have a number of expositional suggestions that seem quite reasonable. In particular, all the referees think that the paper could be written as a more positive contribution. For example, Referee 3 suggests deleting most of the material in Section 2 and Section 5.4, which seems good. The other expositional suggestions of the referees seem good also.*

We deleted Section 2, Section 4 and Section 5.4 in the previous version of the paper.

We thank you and the referees again for giving us an opportunity to improve our paper. I sincerely hope that we have addressed your and the referees' concerns and suggestions satisfactorily in this revised version.

Best Regards,

Che-Lin Su
Assistant Professor
The University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637