Calibrated Projection in MATLAB: Users' Manual*

Hiroaki Kaido†   Francesca Molinari‡   Jörg Stoye§   Matthew Thirkettle¶

October 24, 2017

Abstract

We present the calibrated-projection MATLAB package implementing the method to construct confidence intervals proposed by Kaido, Molinari, and Stoye (2017). This manual provides details on how to use the package for inference on projections of partially identified parameters. It also explains how to use the MATLAB functions we developed to compute confidence intervals on solutions of nonlinear optimization problems with estimated constraints.

Keywords: Partial identification; Inference on projections; Moment inequalities; Uniform inference.

1   Introduction

This manual details the structure of the Calibrated Projection Interval (CPI) algorithm and `MATLAB` Package. It accompanies the paper "Confidence Intervals for Projections of Partially Identified Parameters" (Kaido et al., 2017) and it assumes familiarity with that paper.[1] The CPI algorithm uses an EAM (evaluate, approximate, maximize) algorithm to solve:

$$\inf / \sup_{\theta \in \Theta} p'\theta$$

$$\text{s.t.} \sqrt{n} \frac{\bar{m}_j(\theta)}{\hat{\sigma}_j(\theta)} \leq \hat{c}(\theta) \quad j = 1, \cdots, J,$$

where $\hat{c}(\theta)$ is the calibrated critical value (Jones, Schonlau, & Welch, 1998; Jones, 2001). This version of the CPI algorithm is optimized for basis projection $p = (0, \cdots, 0, 1, 0, \cdots, 0)$ with hyperrectangle parameter constraints $\Theta = \{\theta \in \mathbb{R}^d : \theta_{LB} \leq \theta \leq \theta_{UB}\}$. We also allow for p to be in the unit sphere and polytope constraints on the parameter space, so that $\Theta = \{\theta \in \mathbb{R}^d : \theta_{LB} \leq \theta \leq \theta_{UB}, A_\theta \theta \leq b_\theta\}$. Additional care is required within these extensions (see Appendix B for further details). The current version of the package is written for moment (in)equalities that are separable in data W and parameter θ, so that $E_P[m_j(W_i, \theta)] = E_P[f_j(W_i)] + g_j(\theta)$.[2] Future releases of the package will include:

- Non-separability of $E_P[m_j(W_i, \theta)]$ in W_i and θ.

- Objective function $h(\theta)$ not necessarily equal to $p'\theta$.

We have structured the code so that it is portable. In order to implement a user-specified model, the user needs only input the data, algorithm options, the function that defines the estimators for the moment (in)equalities, as well as the gradients and standard deviation

---

[1]Some notation differs between this paper and (Kaido et al., 2017). This is made clear throughout this manual. Unless otherwise specified, we use notation from the earlier version of the paper (Kaido, Molinari, & Stoye, 2016). The table numbering references (Kaido et al., 2017).

[2]In this manual and in the CPI MATLAB package data is defined as W. The function f and g are the two components of the separable moment (in)equality $E_P[m_j(W_i, \theta)]$. This is in contrast to Kaido et al. (2017), where data is X, $f(\theta)$ refers to the objective function, and $\bar{g}(\theta)$ appears in the EAM algorithm. The subscript n has also been dropped from all estimators.

estimators of the moment functions. Section 2 details how to use the portable code with a user-specified partially identified model with separable moment inequalities. Section 2 also provides instructions on how to replicate the simulations in Kaido et al. (2017). Section 3 provides a deeper insight into how the CPI algorithm is structured. Section 4 discusses extensions to the algorithm and concludes.

2 Using the Calibrated Projection Interval Algorithm

In this section we detail the steps required to run the simulations in Kaido et al. (2017), and how to run a user-specified model. We use as a working example the Entry Game Model in Section 5 of Kaido et al. (2017). This section is organized as follows. Section 2.1 briefly describes the key files in the package. Section 2.2 details how to set up `CVXGEN` and `CVX`, both are fast disciplined convex solvers that we use to compute the calibrated critical value $\hat{c}(\theta)$ (Mattingley & Boyd, 2012; Grant & Boyd, 2014, 2008). Section 2.3 provides instructions on how to replicate the simulations to reproduce the tables in Kaido et al. (2017). Section 2.4 provides instructions on how to implement a user-specified model and compute Projection Intervals (either Calibrated or Andrew and Soares (AS) (Andrews & Soares, 2010)) using the CPI algorithm.

2.1 Overview of Important Files and Folders

First, we briefly describe the key `MATLAB` files and folders.

- `KMS_Simulation.m`. This executes the simulations in Kaido et al. (2017). The DGP, method (Calibrated Projection, Andrew and Soares (AS), or Bugni, Canay, and Shi (BCS)-Profiling),[3] nominal significance level, projection directional vector, number of observations, and number of simulations are set by the user here. The data is generated

---

[3]The code implementing BCS is the code provided by these authors and is available at `http://qeconomics.org/ojs/index.php/qe/article/view/431`.

and passed to either `KMS_0_Main.m` or `BCS_Main`, which computes the Calibrated or AS Projection Interval, or the BCS-Profiled Interval, respectively.

- `KMS_0_Main.m`. This is the file that the user calls to execute the CPI algorithm and compute the Projection Interval (either Calibrated or AS). The user specifies data `W`, the initial guess for a feasible parameter `theta_0`, the projection direction `p`, a set of pre-specified feasible points `theta_feas`, the lower bound on parameter space `LB_theta`, the upper bound on parameter space `UB_theta`, the polytope constraints on the parameter space `A_theta` and `b_theta` so that $A_\theta \theta \le b_\theta$, the nominal significance level `alpha`, a one-sided or two-sided confidence interval `type`, the projection method (calibrated or AS) `CI_method`, the GMS tuning parameter `kappa`, the GMS function `phi`, the name of the `MEX` files for `CVXGEN` (discussed in Section 2.2 below) `CVXGEN_name`, and a structure of algorithm options `KMSoptions`.

The package assumes that the moment (in)equalities are separable, so that $E_P[m_j(W_i, \theta)] = E_P[f_j(W_i)] + g_j(\theta)$.

- `moments_w.m` is the user-specified function for the estimator of $E_P[f_j(W_i)]$, namely $\hat{f}_j$. We allow for both moment inequalities and equalities, as well as paired moment inequalities. If $f_j(W_i)$ is a Bernoulli random variable and if its expectation is too close to 0 or 1, then the corresponding moment (in)equalities are dropped. The output `f_ineq_keep` and `f_eq_keep` defines the moment (in)equalities that are not discarded.

- `moments_theta.m` is the user-specified function for $g_j(\theta)$.

- `moments_gradient.m` is the user-specified function for the gradient of $g_j(\theta)$, which is denoted $D_\theta g_j(\theta)$.

- `moments_stdev.m` is the user-specified function for the estimator for the standard deviation $\sigma_j(W_i)$.

- `KMSoptions.m` defines a structure of algorithm options. `KMSoptions` is also passed to the four user-specified functions above, so the user can pass additional parameters through `KMSoptions` to the user-specified functions (e.g., the support for data W_i). The function `KMSoptions.m` is called before running `KMS_0_Main.m`, and is passed through the last argument of `KMS_0_Main.m`, which is `KMSoptions`.

- `Rho_Polytope_Box.m` and `bound_transform.m` are additional user-written functions needed when polytope constraints on the parameter space are provided (see the arguments `A_theta` and `b_theta` in `KMS_0_Main.m`) or when p is not a basis vector. If p is a non-basis vector or if polytope constraints on the parameter space are included, then sensitivity in the estimate for the projection interval can arise. Details about these files, as well as more detail on the sensitivity issues, is discussed in Appendix B.

The disciplined convex solver `CVXGEN` is used to check whether the set

$$\Lambda^b(\theta, \rho, c) = \{\lambda \in \sqrt{n}(\Theta - \theta) \cap \rho B^d : \mathbb{G}_j^b + D_\theta g_j(\theta)\lambda + \varphi_j(\hat{\xi}_j(\theta)) \leq c, j = 1, \cdots, J\}$$

is empty for each bootstrap repetition $b = 1, \cdots, B$. In order to run `CVXGEN`, the user first compiles a `MEX` file that defines the parameters of the problem (details in Section 2.2).

- The compiled `MEX` files are stored in the subfolder `\CVXGEN`. The file name for this is chosen by the user. For example, we choose `csolve_DGP8.mex64` for the BCS Entry Game. The file name must also be defined when `KMS_0_Main.m` is called. The name is passed via the argument `CVXGEN_name`.

2.2 CVXGEN and CVX Setup

The calibrated critical value $\hat{c}(\theta)$ is computed using a fixed-point algorithm. The fixed-point mapping is computed by checking whether the following set is empty:

$$\Lambda^b(\theta, \rho, c) = \{\lambda \in \sqrt{n}(\Theta - \theta) \cap \rho B^d : \mathbb{G}_j^b(\theta) + D_\theta g_j(\theta)\lambda + \varphi_j(\hat{\xi}_j(\theta)) \leq c, j = 1, \cdots, J\}. \quad (1)$$

This amounts to solving many linear programs (LP), which is done using the fast disciplined convex solver `CVXGEN` (Mattingley & Boyd, 2012) or `CVX` (Grant & Boyd, 2014, 2008).

2.2.1 CVXGEN Setup

To set up `CVXGEN`, the user needs to: 1) install a `MEX` Compiler; 2) generate `C` code at `https://cvxgen.com`; 3) compile and save the `MEX` file; 4) Instruct the CPI algorithm to use `CVXGEN` rather than `CVX`.

The first step is to install a `MEX` compiler. We use the `MinGW-w64 Compiler` on a Windows machine, which is an add-on in `MATLAB`. To install: open `MATLAB`, go to `Home` tab, go to `Add-Ons`. An add-on search window appears on the screen. Search `MinGW-w64 Compiler` and install `MATLAB Support for MinGW-w64 C/C++ Compiler v`. On a Mac, a `C` compiler is supplied with `Xcode`. On a Linux based system, one can use `GCC` (GNU Compiler Collection).

The next step is to generate the `C` code for a specific problem. First, create an account at `https://cvxgen.com` and log in. Next, navigate to the `edit` tab under `problem`. Copy-and-paste the following:

```
dimensions
     dim_p       = XX
     J1          = YY
     J2          = ZZ
     S           = VV
end
parameters
```

```
    A    (J1 + 2*J2 + 2*dim_p + 2 + S , dim_p)
    b    (J1 + 2*J2 + 2*dim_p + 2 + S , 1)
end
variables
    x(dim_p,1)
end
minimize
    0
subject to
    A*x<= b
end
```

Replace `XX` with the dimension of the parameter θ, `YY` with the number of moment inequalities, `ZZ` with the number of moment equalities (do not double count $E_P[m_j(W_i, \theta)] \leq 0$ and $-E_P[m_j(W_i, \theta)] \leq 0$ here), and `VV` with the number of polytope box constraints. If no polytope constraints $A_\theta \theta \leq b_\theta$ are included, set `VV`$= 0$.

Next, navigate to the `generate C` tab under `CODEGEN`. Click `Generate code`. As a result, a list of files populate the webpage. Download the `cvxgen.zip` file and extract. Run `make_csolve.m`. The file `csolve.mex64` should appear in the folder (if on a Linux or Mac machine, the extension is slightly different).[4] Rename `csolve.mex64` to `CVXGEN_name.mex64` (where `CVXGEN_name` is specified by the user) and move the file to the subfolder `\CVXGEN`.

Last, set `KMSoptions.CVXGEN = 1` to instruct CPI algorithm to use `CVXGEN`.

There is an upper bound of $4,000$ non-zero Karush-Kuhn-Tucker matrix entries for the linear program in `CVXGEN`. The size of the problem is determined jointly by J_1, J_2, and d. As an example, `CVXGEN` can handle $\theta \in \mathbb{R}^{10}$ with $J_1 = 55$ and $J_2 = 55$.

2.2.2   CVX Setup

An alternative solver to `CVXGEN` is `CVX`. This solver is slower than `CVXGEN`, but can handle significantly larger LPs and, in our experience, is significantly faster than `MATLAB`'s LP solver

---

[4]If an error occurs here, it is likely that the `MEX` compiler is not installed correctly.

LINPROG. CVX is a `MATLAB` "wrapper" for five different disciplined convex solvers (Grant & Boyd, 2014, 2008). Among these, the solver `MOSEK` is the fastest for our problem. To run `CVX` with `MOSEK`:

1. Ensure that there is a copy of `CVX` is located in the subfolder `\CVX`. If not, navigate to `http://cvxr.com/cvx/` and deposit a copy in the subfolder `\CVX`.

2. Request a license from `http://cvxr.com/cvx/` and deposit it in the same folder.

3. Run `cvx_setup.m`.

4. Set solver using the command `cvx_solver MOSEK` in the `MATLAB` command window.

5. Set `KMSoptions.CVXGEN = 0`.

6. Set `CVXGEN_name` to the empty set.

Once `CVXGEN` or `CVX` is set up, either a simulation model (Section 2.3) or a user-specified model (Section 2.4) can be called via the CPI algorithm.[5]

2.3 Running Simulations

In this section we discuss how to replicate the simulation results in Kaido et al. (2017) (see Tables 1-7 in the paper). As per `CVXGEN` policy, we are unable to distribute the `MEX` files for these simulations. So the first step is to generate the relevant `MEX` files, see Section 2.2 for instructions and Table 1 for `CVXGEN` parameters and naming conventions.

The next step is to set parameters in `KMS_Simulation.m`. Open an instance of `KMS_Simulation.m` and set the following:

- `method = 'KMS'` to compute the Calibrated Projection Interval; or `method = 'AS'` to compute the AS Projection Interval.

---

[5]For additional help with `CVXGEN` or `CVX`, please visit `https://cvxgen.com` and `http://cvxr.com/cvx/`.

- `DGP=k` where k$\in \{1, 2, 3, 4, 5, 6, 7, 8\}$. This parameter selects the data-generating process. $k = 1 - 4$ corresponds to the rotated box described in the earlier version Kaido et al. (2016). $k = 5 - 8$ corresponds to the Entry Games: $k = 5$ is the point-identified Entry Game with zero correlation (Table 3); $k = 6$ is the partially-identified Entry Game with zero correlation (Tables 4, 6 and 7); $k = 7$ is the partially-identified Entry Game with $Corr(u_1, u_2) = 0.5$ (Table 5); $k = 8$ is the BCS simulation (Tables 1-2 in Kaido et al. (2017)).

- `KMS=1` or `KMS=0` determines if `KMS_0_Main` or `BCS_Main` is called. `KMS=0` is a valid input only if `DGP=8`, and `component=1` or `component=2`.

- `component=k` where k$\in \{1, \cdots, \text{dim\_p}\}$ selects the projection direction. That is, the projection vector is p with $p_i = 1$ if $i = k$ and $p_i = 0$ otherwise.

- `n` is the sample size. `n` is set to 4000 for Tables 1-7.

- `Nmc` is the number of Monte Carlo simulations requested. `Nmc` is set to 300 in Table 1 and 1000 in Tables 2-7.

- `sim_lo` and `sim_hi` determine which simulations are run. These parameters are used to split the simulations into batches if needed.

Among other things, convergence criteria are set in `KMSoptions`. All DGPs other than the correlated error DGP, which is DGP 7, use what we call the baseline options. DGP 7, on the other hand, is a fairly difficult problem to solve. Therefore, we use more stringent convergence criteria for this DGP. The options listed below and, in particular, the contraction rates are discussed in more detail in Section 3.2.

The baseline options are:

- `KMSoptions.EAM_maxit=20`. This sets the maximum number of EAM iterations to 20.

- `KMSoptions.h_rate=1.8`. This determines the contraction rate of the parameter space for the M-step.

- `KMSoptions.h_rate2=1.25`. This determines the contraction rate of the parameter space for additional points

- `KMSoptions.EAM_obj_tol = 0.005`. One requirement for convergence is that the absolute difference between the expected improvement projection and the current feasible point $\theta^{*,L}$ is less than `EAM_obj_tol`.

- `KMSoptions.EI_points=10`. The M step is initialized with a set of starting points. The algorithm selects `EI_points` points around the current feasible point $\theta^{*,L}$ that have positive expected improvement. Additional points are also selected.

The stringent options for `DGP=7` are:

- `KMSoptions.EAM_maxit=50`.

- `KMSoptions.h_rate=1.25`.

- `KMSoptions.h_rate2=1.15`.

- `KMSoptions.EAM_obj_tol = 0.0001`.

- `KMSoptions.EI_points=20`.

The number of bootstrap repetitions is also set in `KMSoptions.m`. Table 1 sets this number equal to 301, so that `KMSoptions.B=301`. For Tables 2-7 set `KMSoptions.B=1001`.

Finally, run `KMS_Simulation` to run a simulation with the parameters and options specified above. The results are saved in the subfolder `\Results`.

The file `Analysis.m` carries out post analysis for a particular set of simulations. To run the post analysis, load a results file and run `Analysis.m`. The output includes the median lower bound for the Calibrated Projection Interval; the median upper bound for the Calibrated Projection Interval; coverage percent at the end points of the identification region, as well as at the true parameter; average $\hat{c}(\theta)$; and average computational time.

2.4 User-specified Model

In this section we detail the files that need to be modified so that a user can implement the algorithm to compute the Calibrated Projection Interval for a user-specified problem. We use `DGP=6` as a working example (this is the partially-identified Entry Game with zero correlation between the unobservable shocks u_1 and u_2). For tractability, we have created a working example file `ExampleDGP6.m`. Here, we generate one data set with $n = 4000$ observations. We assume that the parameter space is a hyperrectangle and p is a basis vector throughout the rest of this section.

Step 1: Set up `CVXGEN` or `CVX`

The first step is to set up either `CVXGEN` or `CVX`. This is described in detail in Section 2.2. The required parameters are the dimension of the parameter θ, `dim_p`, the number of moment inequalities, `J1`, and the number of moment equalities, `J2`. In our working example, set `dim_p = 8`, `J1 = 8`, and `J2 = 8`. Name the `MEX` file as `exampleDGP6.mex64` and deposit it in the subfolder `\CVXGEN`. Set `KMSoptions.CVXGEN = 1`. Alternatively, if `CVX` is used, follow the installation instructions at the end of Section 2.2 and set `KMSoptions.CVXGEN = 0`.

Step 2: Moment (In)equality: Estimator for $E_P[f_j(W_i)]$

The moment (in)equalities are separable so that $E_P[m_j(W_i,\theta)] = E_P[f_j(W_i)] + g_j(\theta)$ for $j = 1, \cdots , J$. The estimator for $E_P[f_j(W_i)]$ is defined in the function `moments_w.m`. Override the file `moments_w.m` with the following shell:

```
function [f_ineq,f_eq,f_ineq_keep,f_eq_keep,paired_mom,J1,J2,J3] ...
            = moments_w(W,KMSoptions)
    f_keep_threshold  = KMSoptions.f_keep_threshold;
    J1                = XX1;
    J2                = XX2;
    J3                = XX3;
    f_ineq            = zeros(J1,1);
```

```
    f_eq                 = zeros(2*J2,1);

    f_ineq_keep          = zeros(J1,1);

    f_eq_keep            = zeros(2*J2,1);

    paired_mom           = zeros(J3,1);
%% Define  output  here %%

end
```

The inputs are data W, which is n-by-d_W and `KMSoptions`. The outputs are:

1. `f_ineq` is a J_1-by-1 vector of moment inequalities. As an example, consider `Lines 208-212` in `moments_w.m`, which correspond to the Entry Game moment inequalities (See Equations 5.3-5.4 in Kaido et al. (2017)):

$$E_P[m_j(W_i, \theta)] = E_P[f_j(W_i)] + g_j(\theta), \quad \text{with}$$

$$E_P[f_j(W_i)] = E_P[\mathbb{1}(Y = (0,1))\mathbb{1}(X = x)] \quad \text{and}$$

$$g_j(\theta) = -G_r((-\infty, -x_1'(\beta_1 + \Delta_1) \times [-x_2'\beta_2, \infty)))p_x$$

$$E_P[m_{j+1}(W_i, \theta)] = E_P[f_{j+1}(W_i)] + g_{j+1}(\theta), \quad \text{with}$$

$$E_P[f_{j+1}(W_i)] = -E_P[\mathbb{1}(Y = (0,1))\mathbb{1}(X = x)] \quad \text{and}$$

$$g_{j+1}(\theta) = [G_r((-\infty, -x_1'(\beta_1 + \Delta_1) \times [-x_2'\beta_2, \infty))$$
$$- G_r((-x_1'\beta_1, -x_1'(\beta_1 + \Delta_1) \times [-x_2'\beta_2, -x_2'(\beta_2 + \delta_2)))]p_x,$$

where $W = (Y, X)$.[6] Moment inequalities $j \in \{1, 3, 5, 7\}$ and $j + 1$ correspond to a point x in the support $\mathcal{X} \equiv \{(-1, -1), (-1, 1), (1, -1), (-1, -1)\}$. For each $x \in \mathcal{X}$, $E_P[f_{j+k}(W_i)]$, $k = 0, 1$, is estimated by:

$$\hat{f}_j = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(Y_i = (0,1))\mathbb{1}(X_i = x)$$

$$\hat{f}_{j+1} = -\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(Y_i = (0,1))\mathbb{1}(X_i = x).$$

This corresponds to `Lines 208-212` in `moments_w.m`:

```
f_ineq((ii−1)*2 + 1,1) = sum(Y1 == 0 & Y2 == 1  & X1 == x1 & X2 == x2)/n;
```

---

[6]We define data W with firm-entry decisions Y and market characteristics X. Kaido et al. (2017) define data X with firm-entry decisions Y and market characteristics Z.

```
f_ineq((ii-1)*2 + 2,1) = - sum(Y1 == 0 & Y2 == 1  & X1 == x1 & X2 == x2)/n;
```

2. `f_eq` is a $2J_2$-by-1 vector of moment equalities. Entries $j = 1, \cdots, J_2$ of `f_eq` are defined on `Lines 215-218` in `moments_w.m` and correspond to the moment equalities with $E_P[m_j(W, \theta)] \leq 0$. Entries $j = J_2 + 1, \cdots, 2J_2$ of `f_eq` correspond to $-E_P[m_j(W, \theta)] \leq 0$. It is important to include both the positive and negative of $E_P[m_j(X, \theta)]$ for the moment equalities, see `Line 222` in `moments_w.m`: `f_eq = [f_eq ; -f_eq];`.

3. `f_ineq_keep` and `f_eq_keep` are J_1-by-1 and $2J_2$-by-1 vectors of indicators. These define which moment (in)equalities we keep. If $f_j(W)$ has unbounded support, then set the corresponding entry in `f_ineq_keep` and `f_eq_keep` equal to 1 (see the rotated box example on `Lines 87-88` in `moments_w.m`). In the Games Example, $f_j(W_i)$ is bounded by 0 and 1. If $E_P[f_j(W_i)]$ is close to 0 or 1, Assumption 4.1-(iv) in Kaido et al. (2017), which is taken from Andrews and Soares (2010) and common in the literature, is violated. Therefore, if $\hat{f}_j$ is within the tolerance of `KMSoptions.f_keep_threshold` of 0 or 1, then the corresponding component of `f_ineq_keep` or `f_eq_keep` is set equal to 0, indicating that moment is dropped completely from the analysis – it is not used to compute $\hat{c}(\theta)$ and it does not enter the M-step (see Section 3.2, Pg 24).[7] Otherwise it is set to 1. See `Lines 235-236` in `moments_w.m`. `KMSoptions.f_keep_threshold` is a user-specified option with default value equal to 10^{-4} and may be modified in different applications.

4. `paired_mom` is a J_3-by-1 vector indicating the paired moment inequalities. If there are no paired moment inequalities, set `paired_mom` to the empty set. For each paired moment inequalities, set the corresponding elements in `paired_mom` equal to a unique indicator $j = 1, \cdots, J_3$. See `Lines 227-230` in `moments_w.m`.

5. J_1, J_2, and J_3 define the number of moment inequalities, equalities, and paired moment

---

[7]In the Games Example, the moment inequality functions $f_j(W_i)$ for $j \in \{2, 4, 6, 8\}$ are bounded by -1 and 0. Thus moment inequality $j \in \{2, 4, 6, 8\}$ is dropped for the analysis if $\hat{f}_j$ is too close to -1 or 0.

inequalities.

In the shell above, replace `XX1` - `XX3` with the number of moment inequalities, moment equalities, and paired moment inequalities. Preset each output to the zero vector or the empty set as described above. Last, input user-specified functions for each output.

Step 3: Moment (In)equality: Model-implied Moment Function $g_j(\theta)$

The model-implied function $g_j(\theta)$ is defined in the function `moments_theta.m`. Override `moments_theta.m` with the following shell:

```
function [g_ineq,g_eq] = moments_theta(theta,J1,J2,KMSoptions)
    g_ineq                = zeros(J1,1);
    g_eq                  = zeros(2*J2,1);
%% Define output here %%
end
```

The inputs are the d-by-1 parameter vector `theta`, number of moment inequalities `J1`, number of moment equalities `J2`, and algorithm options `KMSoptions`. The outputs are the J_1-by-1 vector of moment inequalities `g_ineq` and the $2J_2$-by-1 vector of moment equalities `g_eq`, where entries $j = 1, \cdots, J_2$ of `g_eq` correspond to $E_P[m_j(W_i, \theta)] \leq 0$ and entries $j = J_2 + 1, \cdots, 2J_2$ of `g_eq` correspond to $-E_P[m_j(W_i, \theta)] \leq 0$.

In the shell above, input user-specified functions for outputs `g_ineq` and `g_eq`.

As an example, consider `DGP=6`. The moment functions are defined on `Lines 100-169` in `moments_theta.m`. For example, the moment inequality in Equation (5.3) in Kaido et al. (2017) is:

$$
\begin{aligned}
E_P[m_j(W_i, \theta)] &= E_P[f_j(W_i)] + g_j(\theta) \\
&= E_P[f_j(W_i)] + \left[-G_r((-\infty, -x_1'(\beta_1 + \Delta_1) \times [-x_2'\beta_2, \infty)))p_x \right],
\end{aligned}
$$

where $x \in \mathcal{X}$ and p_x is the probability of support point x occurring. $G_r(\cdot, \cdot)$ is the Bivariate Gaussian process with correlation r. `DGP=6` assumes $r = 0$, so the moment $g_j(\theta)$ can be

expressed as:

$$g_j(\theta) = -\Phi(-x_1'(\beta_1 + \Delta_1))(1 - \Phi(-x_2'\beta_2))p_x,$$

where $\Phi(\mu)$ is the univariate Gaussian CDF with mean μ and variance equal to 1. Compare to `Line 143` in `moments_theta.m`:

```
g_ineq((ii-1)*2 + 1,1) = normcdf(-x1*(beta1+delta1))*(1-normcdf(-x2*beta2))*pX;
```

(The negative of `g_ineq` is reported on `Line 164` in `moments_theta.m` to get the correct sign.)

Step 4: Standard Deviation Estimator for $\sigma_j(W_i)$

Under the assumption that the moment functions are separable, the standard deviation does not depend on θ. Specify the standard deviation estimator in the function `moments_stdev`. Override `moments_stdev.m` with the following shell:

```
[f_stdev_ineq , f_stdev_eq] = moments_stdev (W, f_ineq , f_eq , J1 , J2 , KMSoptions)
    f_stdev_ineq          = zeros (J1 ,1);
    f_stdev_eq            = zeros (2*J2 ,1);
%% Define output here %%
end
```

The inputs are: data `W`, data-implied moment functions `f_ineq` and `f_eq`, number of moment (in)equalities `J1` and `J2`, and a structure of options `KMSoptions`. The outputs are the J_1-by-1 vector of standard deviations for the moment inequalities `f_stdev_ineq` and the $2J_2$-by-1 vector of standard deviations for the moment equalities `f_stdev_eq`.

For `DGP=6`, the estimator for $\sigma_j(W_i)$ is

$$\hat{\sigma}_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}(Y_i = y, X_i = x)\left(1 - \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}(Y_i = y_i, X = x)\right)}$$

Compare to `Lines 58–59` in `moments_stdev.m`:

```
f_stdev_ineq (:,1) = sqrt (abs (f_ineq).*(1-abs(f_ineq)));
f_stdev_eq (:,1)   = sqrt (abs (f_eq).*(1-abs(f_eq)));
```

Step 5: Gradient of Model-implied Moment Function $Dg_j(\theta)$

The CPI algorithm requires that the user specifies gradients of the moment functions. Since the moments are separable, the gradient does not depend on data W, so that:

$$\frac{\partial m_j(W, \theta)}{\partial \theta_k} = \frac{\partial g_j(\theta)}{\partial \theta_k}.$$

The gradients are specified in the function `moments_gradient.m`, and below is a shell for this function:

```
[Dg_ineq, Dg_eq] = moments_gradient(theta, J1, J2, KMSoptions)
    dim_p           = KMSoptions.dim_p;
    Dg_ineq         = zeros(J1, dim_p);
    Dg_eq           = zeros(2*J2, dim_p);
%% Define output here %%
end
```

The inputs are: a d-by-1 parameter vector `theta`, the number of moment (in)equalities `J1` and `J2`, and the structure of options `KMSoptions`. The outputs are the J_1-by-d matrix of gradients for the moment inequalities `Dg_ineq`, where

$$\texttt{Dg\_ineq}_{j,k} = \frac{\partial g_j(\theta)}{\partial \theta_k} \quad j = 1, \cdots, J_1, \quad k = 1, \cdots, d$$

and the $2J_2$-by-d matrix of gradients for the moment equalities `Dg_eq`, where

$$\texttt{Dg\_eq}_{j,k} = \frac{\partial g_j(\theta)}{\partial \theta_k} \quad j = J_1 + 1, \cdots, J.$$

As an example, consider the moment inequality in Equation (5.3) Kaido et al. (2017):

$$g_j(\theta) = -\Phi(-x_1'(\beta_1 + \Delta_1))(1 - \Phi(-x_2'\beta_2))p_x.$$

15

The gradients are:

$$\frac{\partial g_j(\theta)}{\partial \beta_1} = x_1' \phi(-x_1'(\beta_1 + \Delta_1))(1 - \Phi(-x_2'\beta_2))p_x$$

$$\frac{\partial g_j(\theta)}{\partial \beta_2} = -x_2' \Phi(-x_1'(\beta_1 + \Delta_1))\phi(-x_2'\beta_2)p_x$$

$$\frac{\partial g_j(\theta)}{\partial \Delta_1} = x_1' \phi(-x_1'(\beta_1 + \Delta_1))(1 - \Phi(-x_2'\beta_2))p_x$$

$$\frac{\partial g_j(\theta)}{\partial \Delta_2} = 0,$$

where $\phi(\mu)$ is the univariate Gaussian PDF with mean μ and variance equal to 1. Compare to `Lines 112-115` in `moments_gradient.m`:

```
Dg3b1 = x1.*normpdf(-x1*(beta1+delta1)) .*(1-normcdf(-x2*beta2))*pX;

Dg3b2 = -x2.*normcdf(-x1*(beta1+delta1)).*normpdf(-x2*beta2)*pX;

Dg3d1 = x1.*normpdf(-x1*(beta1+delta1)) .*(1-normcdf(-x2*beta2))*pX;

Dg3d2 = zeros(1,2);
```

Step 6: Algorithm Options

Algorithm options are specified in the file `KMSoptions.m`. These options should be adjusted for each user-specified model in order to balance computational time and accuracy. The key options are highlighted below.

- `KMSoptions.parallel` turns on parallel computing if set equal to 1 (`Line 35`).

- `KMSoptions.CVXGEN` uses `CVXGEN` if set equal to 1 (`Line 38`).

- `KMSoptions.B` specifies the number of bootstrap repetitions (`Line 42`).

- `KMSoptions.EAM_maxit` specifies the maximum number of EAM iterations (`Line 43`).

- `KMSoptions.mbase` sets the base-multiplier for the initial number of points in the EAM algorithm (`Line 44`). In order to get a better initial approximating surface increase this number. There is a trade off between allowing the EAM algorithm to better approximate the surface near the global maximizer and obtaining a good initial fit.

- `KMSoptions.h_rate` determines the rate at which the parameter space is contracted (`Line 45`). Set equal to a number between 1 and 2. See the M-step in Section 3.2 on Page 25.

- `KMSoptions.h_rate2` should be set to a number between 1 and `KMSoptions.h_rate` (`Line 46`). See the Section 3.2.

- `KMSoptions.EAM_obj_tol` is one of the convergence criteria (`Line 48`). It is required that the absolute difference between the expected improvement maximizer value and $p'\theta^{*,L}$ is less than or equal to `KMSoptions.EAM_obj_tol`.

- `KMSoptions.EAM_maxviol_tol` is another convergence criterion (`Line 56`). It demands that the maximum moment violation is close to 0. Set equal to `inf` to turn off.

- `KMSoptions.EI_points` sets the minimum number of initial points with positive expected improvement for the M step `Line 132`.

Step 7: Determine Input Parameters

The following inputs are required:

1. `W` is an n-by-d_W matrix of data. n is the number of observations and d_W is the number of variables in W. In our working example, $n = 4000$ is the number of markets and $d_W = 6$. The first and second variables in W are the entry decisions of firms 1 and 2. The fourth and sixth variables in W are the random market characteristics of firms 1 and 2. The third and fifth variables in W are constants. The market characteristic has support $\{-1, 1\}$. See `Lines 41-111` in `ExampleDGP6.m` for more detail on how the data is generated.

2. `theta_0` is a d-by-1 vector. It is the initial guess for the parameter vector θ. We arbitrarily set `theta_0` to be the midpoint in the hyperrectangle $\{\theta \in \mathbb{R}^d : \theta_{LB} \leq \theta \leq$

17

$\theta_{UB}\}$ in the working example. See `Line 29` in `ExampleDGP6.m`. Valid input is any value in the parameter space.

3. `p` is the d-by-1 directional vector in the problem $\sup_{\theta \in \Theta} p'\theta$ subject to $\sqrt{n}\frac{\bar{m}_j(\theta)}{\sigma_j} \leq \hat{c}(\theta)$. Valid input is $p_i = 1$ for any one component and 0 otherwise. See `Lines 30-31` in `ExampleDGP6.m`. See Appendix B for non-basis directional vectors.

4. `theta_feas` is a K-by-d matrix of K feasible θ stacked in row format:

$$\begin{bmatrix} \theta_1' \\ \vdots \\ \theta_K' \end{bmatrix}.$$

A feasible θ is one that satisfies $\sqrt{n}\frac{\bar{m}_j(\theta)}{\sigma_j} \leq \hat{c}(\theta), \forall j = 1, \cdots, J$. This input is optional, and if set to the empty set, the algorithm attempts to find a feasible point using an auxiliary search. `theta_feas` is set equal to `[]` on `Line 115` in `ExampleDGP6.m`.

5. `LB_theta` and `UB_theta` are d-by-1 vectors defining the lower and upper bounds of the parameter space. Valid input is $\theta_{LB} \leq \theta_{UB}$. See `Lines 27-28` in `ExampleDGP6.m`.

6. `A_theta` and `b_theta` are L-by-d and L-by-1 matrices defining the polytope constraints on the parameter space. It is required that the parameter space has a non-empty interior. `A_theta` and `b_theta` are set to `[]`, see `Line 115` in `ExampleDGP6.m`.

7. `alpha` is the nominal significance level. Valid input is a number in $[0, 0.5]$. It is set equal to 0.05, see `Line 6` in `ExampleDGP6.m`.

8. `type` determines if either a two-sided or one-sided confidence interval is computed. Valid input is either `'two-sided'` or `'one-sided'`. `type` is set equal to `'two-sided'` for a two-sided confidence interval, see `Line 21` in `ExampleDGP6.m`.

9. `method` determines whether a Calibrated Projection Interval or an AS Projection Interval is computed. Valid input is either `'KMS'` or `'AS'`. `method` is set equal to `'KMS'`, see `Line 4` in `ExampleDGP6.m`.

10. `kappa` specifies the tuning parameter κ. Valid input is either `NaN` for the default $\kappa = \sqrt{\ln(n)}$ or a user-specified function `@(n)kappa_function(n)` satisfying Assumption 4.2 in Kaido et al. (2017). `kappa` is set equal to `NaN`, see `Line 22` in `ExampleDGP6.m`.

11. `phi` specifies the GMS function $\varphi_j(x)$. Valid input is either `NaN` for the default hard thresholding function

$$\varphi_j(x) = \begin{cases} 0 & \text{if } x \geq -1 \\ -\infty & \text{else} \end{cases}$$

or a user-specified function `@(x)phi_function` satisfying Assumption 4.2. `phi` is set equal to `NaN`, see `Line 23` in `ExampleDGP6.m`.

12. `CVXGEN_name` is the name for the `CVXGEN MEX` file, see Section 2.2. `CVXGEN_name` is set equal to `'ExampleDGP6'`, see `Line 39` in `ExampleDGP6.m`.

13. `KMSoptions` is the structure of algorithm options. It is called on `Line 9` in `ExampleDGP6.m` and updated further throughout the file.

Finally, the CPI algorithm is called on `Lines 114-115` in `ExampleDGP6.m`, with the inputs specified below:

```
[KMS_confidence_interval ,KMS_output] = KMS_0_Main(W, theta_0 , p ,[] , LB_theta ,...
    UB_theta ,[] ,[] , alpha , type , method , kappa , phi ,CVXGEN_name, KMSoptions );
```

3   Calibrated Projection Interval Algorithm

In this section we provide an overview of the CPI algorithm. We start with `KMS_0_Main`, since this is the function that calls the CPI algorithm. The empirical moments $\hat{f}_j$, standard

deviation $\hat{\sigma}_j$, and the recentered bootstrap moments $\mathbb{G}_j^b$ do not depend on θ and can be computed outside of the EAM algorithm. These are computed on `Lines 217, 502`, and `Lines 438-493` and `509-510` in `KMS_0_Main`, respectively. The recentered bootstrap moments are denoted `G_ineq` and `G_eq`.

It is required that the EAM algorithm is initiated with a feasible point, that is, a point $\theta^{\text{feas}} \in \Theta$ satisfying:

$$\sqrt{n}\frac{\hat{f}_j + g_j(\theta^{\text{feas}})}{\hat{\sigma}_j} \leq \hat{c}(\theta^{\text{feas}}) \quad \forall j = 1, \cdots, J.$$

The CPI algorithm executes two feasible search algorithms on `Lines 565-583` in `KMS_0_Main`. The feasible search algorithms are `KMS_1_FeasibleSearch.m` and `KMS_2_EAM_FeasibleSearch.m`. If a feasible point(s) is supplied by the user, then the algorithm skips this step.

The EAM algorithm is called on `Lines 602-635` in `KMS_0_Main`. The search direction p is executed first and the search direction $-p$ second. Output including the optimal point $\theta^{*,EAM}$, the calibrated critical value at this point $\hat{c}(\theta^{*,EAM})$, the expected improvement $EI(\theta^{*,EAM})$, and the convergence time is reported in the structure `KMS_output`.

An optional algorithm on `Lines 637-676` in `KMS_0_Main` is also included (set `KMSoptions.direct_solve=1` to run this algorithm). This algorithm solves the problem:

$$\min/\max_{\theta \in \Theta} p'\theta$$
$$\text{s.t. } \sqrt{n}\frac{\hat{f}_j + g_j(\theta)}{\hat{\sigma}_j} \leq \hat{c}(\theta) \quad \forall j = 1, \cdots, J$$

using numerical gradients (the gradient of $\hat{c}(\theta)$ is unknown). Even for simple problems this algorithm requires a large amount of computational time to find a solution and the solution is often not the global minimizer/maximizer.[8]

Finally, the Calibrated Projection Interval (or AS Projection Interval) is reported on

---

[8]One could specify an analytical gradient function, where the analytical gradients $D_\theta g_j(\theta)$ are passed and the numerical gradient for $\hat{c}(\theta)$ is computed. This is not done in this version of the CPI algorithm.

`Lines 678-707` in `KMS_0_Main`. The feasible search and EAM algorithm is discussed in Sections 3.1 and 3.2 below. The algorithm that computes $\hat{c}(\theta)$ is discussed in Section 3.3

3.1 Feasible Search Algorithm

The feasible search algorithms `KMS_1_FeasibleSearch` and `KMS_2_EAM_FeasibleSearch` attempt to find a point θ satisfying:

$$\theta \in \Theta^{\text{feas}} \equiv \left\{ \theta \in \Theta : \sqrt{n} \frac{\hat{f}_j + g_j(\theta^{\text{feas}})}{\hat{\sigma}_j} \leq \hat{c}(\theta^{\text{feas}}) \quad \forall j = 1, \cdots, J \right\}. \tag{2}$$

The algorithm `KMS_1_FeasibleSearch` solves the problem:

$$\min_{\theta \in \Theta} \max_{j=1,\cdots,J} \sqrt{n} \frac{\hat{f}_j + g(\theta)}{\hat{\sigma}_j}. \tag{3}$$

Let the minimizer be $\theta^{*,FS1}$. The hope is that $\theta^{*,FS1}$ satisfies the relaxed condition

$$\sqrt{n} \frac{\hat{f}_j + g_j(\theta^{*,FS1})}{\hat{\sigma}_j} \leq \hat{c}(\theta^{*,FS1}), \forall j = 1, \cdots, J. \tag{4}$$

If $\max_j \sqrt{n} \frac{\hat{f}_j + g_j(\theta^{*,FS1})}{\hat{\sigma}_j} \leq 0$, then the condition in Equation (4) is satisfied since $\hat{c}(\theta) \geq 0, \forall \theta \in \Theta$.

A MultiStart algorithm is used to solve Problem (3).[9] On `Lines 72-78` in `KMS_1_Feasible-Search`, a set of starting points is drawn uniformly from Θ. These starting points are passed to `fmincon` (`Lines 93-119`). At each solution, the constraint violation is computed on `Line 134`. If there is a feasible point (constraint violation = 0), then the feasible point `theta_feas` is returned. Otherwise, `flag_feas=0` is returned indicating failure to find a feasible point.

If the feasible search algorithm `KMS_1_FeasibleSearch` fails, the second feasible search algorithm, `KMS_2_EAM_FeasibleSearch`, is executed. It uses an EAM-type algorithm to try

---

[9]The solver `fmincon` cannot efficiently solve Problem (3) as written, since the gradient of $\max_{j=1,\cdots,J} \sqrt{n} \frac{\hat{f}_j + g(\theta)}{\hat{\sigma}_j}$ is unknown. By introducing a free parameter γ, the problem can be re-written so that `fmincon` can solve it. See Section 3.2.

to find a feasible point. In particular, `Line 86-92` in `KMS_2_EAM_FeasibleSearch.m` draws an initial set of points $\theta^{(1)}, \cdots, \theta^{(L)}$. The calibrated critical value $\hat{c}(\theta)$ is computed at each of these points (E-step, `Line 109`). If any of these points are feasible, the algorithm is terminated and a feasible point is returned (`Line 125-137`). Otherwise, the surface $c_L(\theta)$ is approximated using the kriging method (A-step, `Lines 139-144`). Last, the following minimization problem is solved using a MultiStart algorithm:

$$\min_{\theta \in \Theta} \max_{j=1,\cdots,J} \left(\sqrt{n} \frac{\hat{f}_j + g(\theta)}{\hat{\sigma}_j} - c_L(\theta) \right), \tag{5}$$

(M-step, `Lines 174-204`). Call the minimizer $\theta^{*,L+1}$. This EAM algorithm is re-iterated with the new set of points $\{\theta^{(l)}\}_{l=1}^{L} \cup \{\theta^{*,L+1}\}$, and continues until either a feasible point is obtained or the maximum number of iterations `KMSoptions.EAM_maxit` is reached.

3.2   EAM Algorithm

The EAM algorithm (`KMS_3_EAM`) for search direction $q \in \{-p, p\}$ is called on `Lines 603-635` in `KMS_0_Main.m`. The inputs for the EAM algorithm are:

1. `q` is the directional vector, set equal to either p or $-p$.

2. `sgn_q` is equal to -1 or 1. It specifies whether we are maximizing in direction $q = p$ or $q = -p$.

3. `theta_feas` is a K-by-d matrix of feasible points stacked in row format.

4. `theta_init`, `c_init`, `CV_init`, `maxviol_init` are a set of θs, calibrated critical values, constraint violations, and maximum violations passed from the feasible search algorithm. These can be empty.

5. `f_ineq`, `f_eq`, `f_ineq_keep`, `f_eq_keep` are output from `moments_w.m`, which are the data-implied moment (in)equalities and the moment (in)equalities that we keep.

6. `f_stdev_ineq` and `f_stdev_eq` are output from `moments_stdev.m`, which are the standard deviations of the moment (in)equalities.

7. `G_ineq` and `G_eq` are the recentered bootstrap moment (in)equalities.

8. `KMSoptions` is a structure of algorithm options.

The outputs are:

1. `theta_hat` is the d-by-1 solution to $\max_{\theta \in \Theta} q'\theta$ subject to the calibrated moment inequality constraints from the EAM algorithm. Call this optimal point $\theta^{*,EAM}$.

2. `theta_optbound` is the value $q'\theta^{*,EAM}$.

3. `c` is the calibrated critical value $\hat{c}(\theta^{*,EAM})$.

4. `CV` is the maximum constraint violation at $\theta^{*,EAM}$. This should be zero if $\theta^{*,EAM}$ is feasible.

5. `EI` is the expected improvement at $\theta^{*,EAM}$.

6. `flag_opt` is a flag equal to 1 if the EAM algorithm converged.

The key steps in the EAM algorithm are detailed below.

Initialization: A initial set of points is drawn on `Lines 93–99` in `KMS_3_EAM`. The set of feasible points `theta_feas` is also added to the pool of initial points $\{\theta^{(l)}\}_{l=1}^{L_0}$. The set of points to be evaluated is saved in the L-by-d matrix `theta_Estep` (where $L = L_0$ on the first iteration).

The following steps are iterated until either the program converges or a preset maximum number of iterations is reached.

E-step, Evaluation: `Line 143` in `KMS_3_EAM` calls the function `KMS_31_Estep`. Within this function, `c_Estep` is the L-by-1 vector of calibrated (or AS) critical values $\hat{c}(\theta)$ for each

$\theta \in \{\theta^{(l)}\}_{l=1}^{L}$; `CV_Estep` and `maxviol_Estep` are the constraint violation and maximum violation for each $\theta \in \{\theta^{(l)}\}_{l=1}^{L}$, respectively. The subfunctions `KMS_31_Estep`, `KMS_32_Critval`, and `KMS_33_Coverage` are discussed in more detail in Section 3.3.

Lines 145–155 in `KMS_3_EAM` prepare the matrices `theta_Astep` and `c_Astep`, which keep track of all points to be passed to the A-step. The kriging model is sensitive if two points $\theta^{(l)}$ and $\theta^{(k)}$ are too close together. Therefore, if two points are too close to one another, only one point is passed to the kriging model, see Line 161 in `KMS_3_EAM`.

A-step, Approximation: The surface $\hat{c}(\theta)$ is approximated via a kriging model. The set of points $\{\theta^{(l)}, \hat{c}(\theta^{(l)})\}_{l=1}^{L}$ is passed to the kriging function on Line 163 in `KMS_3_EAM`. We use the `DACE` package (Lophaven, Nielsen, & Sondergaard, 2002). The `DACE MATLAB` files are saved in the subfolder `\dace`. The output is the structure `dmodel`. The function `[c,Dc,mse,Dmse]=predictor(theta,dmodel)` uses the interpolated surface to predict the value of $\hat{c}(\theta)$ and gradient $D_\theta \hat{c}(\theta)$ at θ. The standard deviation $\hat{\zeta} s_L(\theta)$ is also estimated and is equal to $\sqrt{\texttt{mse}}$.[10]

M-step, Maximization: Using the approximated surface $c_L(\theta)$ and standard deviation $\hat{\zeta} s_L(\theta)$, the next point in the sequence $\{\theta^{(l)}\}_{l=1}^{L}$ is chosen to maximize the expected improvement function:

$$\theta^{*,\text{Mstep}} = \underset{\theta \in \Theta}{\arg\max}\, \mathbb{EI}(\theta) = (q'\theta - q'\theta^{*,L})_+ \left(1 - \Phi\left(\frac{\max_{j=1,\cdots,J} \hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta} s_L(\theta)}\right)\right), \quad (6)$$

where

$$\theta^{*,L} \equiv \underset{\theta \in \{\theta^{(l)}\}_{l=1}^{L}}{\arg\max}\, q'\theta \quad \text{s.t.} \quad \sqrt{n}\frac{\hat{f}_j + g_j(\theta)}{\hat{\sigma}_j} \le \hat{c}(\theta), \quad \forall j = 1, \cdots, J,$$

$$(x)_+ \equiv \max(0, x),$$

$$\hat{h}_j(\theta) \equiv \sqrt{n}\frac{\hat{f}_j + g_j(\theta)}{\hat{\sigma}_j}.$$

---

[10]For more information and source files for the `DACE MATLAB` package, go to `http://www2.imm.dtu.dk/projects/dace/`.

Problem (6) is solved using a MultiStart algorithm. Three important steps are required to resolve numerical issues with Problem (6):

1. The gradients of the functions $(x)_+$ and $\max_{j=1,\cdots,J} \hat{h}_j(\theta)$ are undefined at crossing points (e.g, at $x = 0$ for $(x)_+$). Problem (6) is rewritten as:

$$- \min_{\theta \in \Theta, q'\theta \geq q'\theta^{*,L}} \max_{j=1,\cdots,J} \left[-(q'\theta - q'\theta^{*,L})\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)} \right) \right].$$

The min / max problem can be solved using `fmincon` by introducing a free parameter γ and rewriting the problem as a minimization problem with nonlinear constraints.

$$\min_{\gamma \in \mathbb{R}, \theta \in \Theta, q'\theta \geq q'\theta^{*,L}} \gamma \tag{7}$$

$$\text{s.t.} \quad -(q'\theta - q'\theta^{*,L})\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)} \right) - \gamma \leq 0, \forall j = 1, \cdots, J.$$

The gradients of the objective function and constraints in Problem (7) are well-defined, so it can be solved using `fmincon` with analytical gradients. See the functions `KMS_34_EI-_objective.m` and `KMS_35_EI_constraint.m`. The same technique is used to solve Problems (3) and (5) in the feasible search algorithms.

2. Problem (6) can become ill-conditioned in the sense that the objective function is numerically equal to zero for θ such that the (in)equalities are modestly violated. Consequently, the non-linear programming solver `fmincon` may get stuck. To overcome this issue, we use an auxiliary method to draw θ such that $\mathbb{EI}(\theta) > 0$ (see `Line 287` in `KMS_3_EAM` and see `KMS_36_drawpoints`). These points are passed to `fmincon` and Problem (7) is solved using MultiStart.

To further explain the numerical issue, observe that the argument of $\Phi(\cdot)$ is

$$-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)}.$$

If the approximated moment condition $\hat{h}_j(\theta) - c_L(\theta)$ is violated and hence positive, the argument of $\Phi(\cdot)$ is negative. If, in addition, $\hat{\zeta}s_L(\theta)$ is small relative to $\hat{h}_j(\theta) - c_L(\theta)$, then the term $\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)}\right)$ can be numerically equal to zero. Therefore, the objective function may have many local minima, and in applications `fmincon` may get stuck.

3. The expected improvement objective function trades off increasing the value of $q'\theta$ and increasing the likelihood that θ satisfies the moment conditions via the term $\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)}\right)$. The approximation $c_L(\cdot)$ to $\hat{c}(\cdot)$ is not perfect. Therefore, the value of $\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)}\right)$ can be positive for $\theta \notin \Theta^{\text{feas}}$ (defined in Equation (2)). In later iterations of the EAM algorithm, we may find that the expected improvement maximizer $\theta^{*,\text{Mstep}}$ is not feasible. Adding non-feasible $\theta^{*,\text{Mstep}}$ improves the fit of $c_L(\theta)$ and $\Phi\left(-\frac{\hat{h}_j(\theta) - c_L(\theta)}{\hat{\zeta}s_L(\theta)}\right)$ converges in probability to 0 for $\theta \notin \Theta^{\text{feas}}$ as the number of EAM iterations increases. To increase the likelihood that $\theta^{*,\text{Mstep}} \in \Theta^{\text{feas}}$, we contract the parameter space, which constrains the expected improvement function. The contraction forces the expected improvement maximizer $\theta^{*,\text{Mstep}}$ to be near $\theta^{*,L}$ if the EAM algorithm begins to stall, increasing the likelihood that $\theta^{*,\text{Mstep}}$ is feasible.

The contracted parameter space is:

$$\Theta(h_{\text{rate}}^{\text{counter}}) = \left\{\theta : \theta_{LB} \leq \theta \leq \theta_{UB}, A_\theta \theta \leq b_\theta, q'\theta^{*,L} \leq q'\theta \leq q'\theta^{*,L} + \frac{q'\theta^\dagger - q'\theta^{*,L}}{h_{\text{rate}}^{\text{counter}}}\right\},$$

$$(8)$$

where $q'\theta^\dagger = \max_{\theta \in \Theta} q'\theta$. For example, if $q = p$, then $q'\theta^\dagger = q'\theta_{UB}$. The term $h_{\text{rate}}^{\text{counter}}$ controls the rate of contraction. h_{rate} is specified by the user and is equal to `KMSoptions.h_rate`. The exponent `counter` is a natural number that increases by one on iteration i of the EAM algorithm if insufficient progress is made on iteration $i-1$. If $\theta^{*,L}$ is too close to the contracted boundary, then `counter` decreases by one. Problem (7) is solved subject to $\theta \in \Theta(h_{\text{rate}}^{\text{counter}})$. Our convergence criteria (described

below) are chosen to make sure that no mechanical convergence occurs.

The contraction of the parameter space occurs on `Lines 200-273` in `KMS_3_EAM`, `Lines 281-328` draw points, and `Lines 336-373` call `fmincon`.

Updating and Convergence: The final step in the EAM algorithm is to update the set of points and check convergence criteria. Provided we find a point with positive expected improvement in the M-step, we add both the M-step solution as well as a uniformly-drawn point to the set of evaluation points $\{\theta^{(l)}\}_{l=1}^{L}$. We also add two points, θ_{ϵ_1} and θ_{ϵ_2} that are close to $\theta^{*,L}$ and satisfy $q'\theta_{\epsilon_k} > q'\theta^{*,L}$. The distance between θ_{ϵ_k} and $\theta^{*,L}$ is determined by the option `KMSoptions.h_rate2`. See `Lines 430-439` in `KMS_3_EAM`.

The convergence check occurs on `Lines 464-485` in `KMS_3_EAM`. We first check if $\theta^{*,L}$ is too close to the boundary of Θ. If $|p'\theta^{*,L} - p'\theta_{UB}| < 10^{-4}$ (for search direction p), then a warning that the parameter is on the boundary is displayed and we output $\theta^{*,\mathrm{EAM}} = \theta^{*,L}$. Otherwise, if all of the following conditions are satisfied, we say that the EAM algorithm has converged, and we output $\theta^{*,\mathrm{EAM}} = \theta^{*,L}$.

1. `iter >= EAM_minit`: The current iteration i of the EAM algorithm is greater than or equal to `KMSoptions.EAM_minit`. This ensures that the EAM algorithm does not terminate early. Default is `KMSoptions.EAM_minit=4`.

2. `change_EI_proj < EAM_obj_tol`: The absolute difference in the value of the objective function $q'\theta$ between the expected improvement maximizer $\theta^{*,\mathrm{Mstep}}$ and the current feasible optimal $\theta^{*,L}$ is less than the tolerance parameter `KMSoptions.EAM_obj_tol`.

3. `change_proj < EAM_tol`: The absolute difference in the value of the objective function $q'\theta$ between the current feasible optimal $\theta^{*,L}$ and the previous iteration's feasible optimal $\theta^{*,L-1}$ is less than the tolerance parameter `KMSoptions.EAM_toll`.

4. `feas_points>num_feas`: We have found at least one feasible point inside the EAM algorithm. This ensures that we do not terminate the EAM algorithm with only the feasible points that are fed into the EAM function.

27

5. `abs(opt_dagger - q.'*theta_hash) > 1e-4`: The point $\theta^{*,L}$ is not too close to the boundary of the contracted parameter space. This ensures that the EAM algorithm is not terminated mechanically.

6. `abs(maxviol_hash) <EAM_maxviol_tol`: The value $\left| \max_{j=1,\cdots,J} \sqrt{n}\frac{\hat{f}_j + g_j(\theta^{*,L})}{\hat{\sigma}_j} - \hat{c}(\theta^{*,L}) \right|$ is less than the tolerance parameter `KMSoptions.EAM_maxviol_to`. If this condition is violated, then by continuity of $\hat{c}(\cdot)$ it is possible to increase the value of $q'\theta^{*,L}$ and not violate the moment conditions.

Last, the contraction counter is updated (`Lines 487-494`). If $|q'\theta^{*,L} - q'\theta^{*,L-1}| < 1 \times 10^{-6}$ (so that no progress is made between this iteration and the previous iteration), then the contraction counter is increased by one: `counter = counter + 1`. If the contraction counter is positive and $\theta^{*,L}$ is too close to the contracted parameter space the contraction counter is decreased by one: `counter = counter - 1`.

3.3   Root-Finding Algorithm Used to Compute $\hat{c}_n(\theta)$

This section explains in detail how to compute the calibrated critical value $\hat{c}(\theta)$:

$$\hat{c}(\theta) \equiv \inf\{c \in \mathbb{R}_+ : P^*(\Lambda^b(\theta,\rho,c) \cap p'\lambda = 0\} \neq \emptyset \geq 1 - \alpha\}, \tag{9}$$

where P^* is the bootstrap empirical frequency. The relevant `MATLAB` files are: `KMS_31_Estep.m`, `KMS_32_Critval.m`, and `KMS_33_Coverage.m`. For a given $\theta \in \Theta$, $P^*(\Lambda^b(\theta,\rho,c) \cap \{p'\lambda = 0\} \neq \emptyset)$ increases in c (with $\Lambda^b(\theta,\rho,c)$ defined in Equation (1)), and so $\hat{c}(\theta)$ can be quickly computed via a root-finding algorithm, such as the Brent-Dekker Method (BDM), see Brent (1971) and Dekker (1969). To do so, define $\Psi_\alpha(c) = \frac{1}{B}\sum_{b=1}^B \psi_b(c) - (1-\alpha)$ where

$$\psi_b(c(\theta)) = \mathbf{1}(\Lambda^b(\theta,\rho,c) \cap \{p'\lambda = 0\} \neq \emptyset).$$

28

Let $\bar{c}(\theta)$ be an upper bound on $\hat{c}(\theta)$, for example, the asymptotic Bonferroni bound $\bar{c}(\theta) \equiv \Phi^{-1}(1 - \alpha/J)$. It remains to find $\hat{c}(\theta)$ such that $\Psi_\alpha(\hat{c}(\theta)) = 0$ when $\Psi_\alpha(0) \leq 0$. It is possible that $\Psi_\alpha(0) > 0$ in which case we output $\hat{c}(\theta) = 0$. Otherwise, we use BDM to find the unique root to $\Psi_\alpha(c)$ on $[0, \bar{c}(\theta)]$ where, by construction, $\Psi_\alpha(\bar{c}(\theta)) \geq 0$. We propose the following algorithm:

Step 0 (Initialize)

1. Set *Tol* equal to a chosen tolerance value;

2. Set $c_L = 0$ and $c_U = \bar{c}(\theta)$ (values of c that bracket the root $\hat{c}_n(\theta)$);

3. Set $c_{-1} = c_L$ and $c_{-2} = \perp$ (proposed values of c from 1 and 2 iterations prior). Also set $c_0 = c_L$ and $c_1 = c_U$.

4. Compute $\varphi_j(\hat{\xi}_j(\theta))$ for $j = 1, \cdots, J$;

5. Compute $D_\theta g_j(\theta)$ for $j = 1, \cdots, J$;

6. Compute $\mathbb{G}_j^b$ for $b = 1, \cdots, B$, $j = 1, \cdots, J$;

7. Compute $\psi_b(c_L)$ and $\psi_b(c_U)$ for $b = 1, \cdots, B$;

8. Compute $\Psi_\alpha(c_L)$ and $\Psi_\alpha(c_U)$.

Step 1 (Method Selection)

Use the BDM rule to select the updated value of c, say c_2. The value is updated using one of three methods: Inverse Quadratic Interpolation, Secant, or Bisection. The selection rule is based on the values of c_i, $i = -2, -1, 0, 1$ and the corresponding function values.

Step 2 (Update Value Function)

Update the value of $\Psi_\alpha(c_2)$. We can exploit the previous computation and monotonicity of the function $\psi_b(c_2)$ to reduce computational time:

1. If $\psi_b(c_L) = \psi_b(c_U) = 0$, then $\psi_b(c_2) = 0$;

2. If $\psi_b(c_L) = \psi_b(c_U) = 1$, then $\psi_b(c_2) = 1$.

Step 3 (Update)

1. If $\Psi_\alpha(c_2) \geq 0$, then set $c_U = c_2$. Otherwise set $c_L = c_2$.

2. Set $c_{-2} = c_{-1}$, $c_{-1} = c_0$, $c_0 = c_L$, and $c_1 = c_U$.

3. Update corresponding function values $\Psi_\alpha(\cdot)$.

Step 4 (Convergence)

1. If $\Psi_\alpha(c_U) \leq$ *Tol* or if $|c_U - c_L| \leq$ *Tol*, then output $\hat{c}_n(\theta) = c_U$ and exit. Note: $\Psi_\alpha(c_U) \geq 0$, so this criterion ensures that we have *at least* $1 - \alpha$ coverage.

2. Otherwise, return to **Step 1**.

The computationally difficult part of the algorithm is computing $\psi_b(\cdot)$ in **Step 2**. This is simplified for two reasons. First, evaluation of $\psi_b(c)$ entails determining whether a constraint set comprised of $J + 2d + 2$ linear inequalities in d variables is feasible. This can be accomplished by efficiently employing `CVXGEN` or `CVX`. Second, we exploit monotonicity in $\psi_b(\cdot)$ to reduce the number of linear programs needed to be solved.

The file `KMS_31_Estep` fixes θ and computes the GMS function $\varphi(\hat{\xi}_j(\theta))$ and gradients $D_\theta g_j(\theta)$; `KMS_32_Critval` executes the BDM algorithm; and `KMS_33_Coverage` computes $\psi_b(c)$ for $b = 1, \cdots, B$.

4 Discussion

We have described how to implement the CPI algorithm to solve

$$\inf / \sup_{\theta \in \Theta} p'\theta \tag{10}$$

$$\text{s.t.} \sqrt{n}\frac{\bar{m}_j(\theta)}{\hat{\sigma}_j(\theta)} \leq \hat{c}(\theta) \quad j = 1, \cdots, J. \tag{11}$$

One difficulty in solving this problem is that $\hat{c}(\theta)$ is a "black-box function" with an unknown gradient. Directly solving this problem using `fmincon` with numerical gradients is slow and can return local solutions that are far from the global solution(s). The EAM algorithm is employed to solve this problem. This manual and `MATLAB` package can serve as a guide on how to implement the EAM algorithm to solve other black-box functions, provided sufficient continuity assumptions hold.

In the next release of the `MATLAB` package, we will allow for non-separability of the moment (in)equalities. Additional numerical issues are presented in this case. The estimator for the moment function, the estimator for the standard deviation, and the bootstrap need to be recomputed at each visit of $\theta \in \Theta$ in the EAM algorithm. The gradients of the moment functions also depend on the data, so this adds additional numerical complexity in the optimization routine.

Another feature we plan to incorporate is an objective function $h(\theta)$ in Equation (10) that is not necessarily equal to $p'\theta$. The objective function, for example, could be a welfare function from a partially identified model that is parameterized by θ (for example, see Barseghyan, Coughlin, Molinari, and Teitelbaum (2017)). Consequently, the welfare function is also partially identified. Using the CPI algorithm one can obtain uniformly valid bounds on the function $h(\theta)$. Modifications to the feasible search, fixed-point algorithm, and M-step are required.

References

Andrews, D. W., & Soares, G. (2010). Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica*, *78*(1), 119–157.

Barseghyan, L., Coughlin, M., Molinari, F., & Teitelbaum, J. (2017). *Limited Consideration and Risk Preferences.* (Work in progress)

Brent, R. P. (1971). An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, *14*(4), 422–425.

Dekker, T. (1969). Finding a zero by means of successive linear interpolation. *Constructive aspects of the fundamental theorem of algebra*, 37–51.

Grant, M., & Boyd, S. (2008). Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, & H. Kimura (Eds.), *Recent advances in learning and control* (pp. 95–110). Springer-Verlag Limited. (`http://stanford.edu/~boyd/graph_dcp.html`)

Grant, M., & Boyd, S. (2014, March). *CVX: Matlab software for disciplined convex programming, version 2.1.* `http://cvxr.com/cvx`.

Jones, D. R. (2001). A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, *21*(4), 345–383.

Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, *13*(4), 455–492.

Kaido, H., Molinari, F., & Stoye, J. (2016). Confidence intervals for projections of partially identified parameters. *arXiv preprint arXiv:1601.00934v1*.

Kaido, H., Molinari, F., & Stoye, J. (2017). Confidence intervals for projections of partially identified parameters. *arXiv preprint arXiv:1601.00934v2*.

Lophaven, S., Nielsen, H., & Sondergaard, J. (2002). Dacea matlab kriging toolbox; version 2; informatics and mathematical modelling. *Technical University of Denmark, Copenhagen*.

Mattingley, J., & Boyd, S. (2012). Cvxgen: A code generator for embedded convex

optimization. *Optimization and Engineering*, *13*(1), 1–27.

Appendices

A Tables

DGP	dim_p	J1	J2	S	CVXGEN_name
1-3	2	4	0	0	csolve_DGP1
4	2	8	0	0	csolve_DGP4
5-6	8	8	8	0	csolve_DGP5
7	9	8	8	0	csolve_DGP7
8	5	8	4	13	csolve_DGP8

Table 1: List of parameters for creating the `CVXGEN MEX` files for simulations in Kaido et al. (2017). The first column corresponds to the parameter `DGP` in `KMS_Simulation`.

B Polytope Constraints and Non-basis Directional Vectors

In this appendix we describe the numerical issues that arise when either p is a non-basis directional vector or polytope constraints are imposed on the parameter space. We also propose a method on how to resolve these issues. The key issue is how to draw points from the contracted parameter space, see Equation (8). If the constraints $A_\theta \theta \leq b_\theta$ are included or if p is not a basis vector, then the contracted parameter space is a polytope but not a hyperrectangle (henceforth, called a non-basis polytope). In either case the numerical problem amounts to drawing points uniformly from a non-basis polytope.

We have identified three methods that can be used to draw points from a non-basis polytope. We, however, find that only the third method is reliable.

1. Hit-and-Run (HR) sampling. HR sampling uses Monte Carlo Markov Chain methods to draw points uniformly from the non-basis polytope $\Theta(h_{\text{rate}}^{\text{counter}}) \subset \mathbb{R}^d$. The method is, however, numerically unstable if the non-basis polytope is thin. The contracted pa-

rameter space in the EAM algorithm converges to a polytope in $\mathbb{R}^{d-1}$ as the contraction counter increases Therefore, HR sampling is unreliable for our problem.

2. Weighted average of vertices. In this method, the vertices of the contracted parameter space $\Theta(h_{\text{rate}}^{\text{counter}})$ are computed. A randomly generated point can be generated from a random weighted average of the vertices. Uniform weights do not guarantee that the point is uniformly drawn from $\Theta(h_{\text{rate}}^{\text{counter}})$. This, never-the-less, does not violate convergence assumptions for the EAM algorithm provided that there is positive mass at all points $\theta \in \Theta(h_{\text{rate}}^{\text{counter}})$. The algorithm that computes the vertices suffers from numerical issues as the parameter space becomes thin, and so this method is not appropriate for the CPI algorithm.

3. Draw-and-Discard sampling (DD). The algorithm first draws points uniformly from a box $B \supset \Theta(h_{\text{rate}}^{\text{counter}})$. It then discards any points that are not in $\Theta(h_{\text{rate}}^{\text{counter}})$. The volume of B relative to $\Theta(h_{\text{rate}}^{\text{counter}})$ must be small for this method to work well. If not, then a large number of initial points are required in order to achieve a target number of points. Therefore, the box B needs to be carefully defined.

In the current version of the CPI algorithm, the DD method only works for when p is a basis vector and the parameter space is a non-basis polytope. Modifications to the user-written function `bound_transform.m` are required. We explain the modifications with an example. The parameter space for DGP 8 is the polytope:

$$\Theta = \{\theta \in \mathbb{R}^5 : \theta_1 \in [0, 1], \theta_2 \in [0, 1], \theta_k \in [0, \min\{\theta_1\, \theta_2\}], k = 3, 4, 5\}.$$

First, to run DD sampling set `KMSoptions.HR=0` (to use hit-and-run sampling set `KMSoptions-.HR=1`). To draw points from this space we use the draw-and-discard sampling method. The file `bound_transform.m` defines the box B above. It is not advised to set B to be the parameter bounds θ_{LB} and θ_{UB}, as the volume of this box relative to the contracted parameter space $\Theta(h_{\text{rate}}^{\text{counter}})$ quickly diverges. The inputs of `bound_transform` are: `LB_in`, `UB_in`, and

KMSoptions. The inputs LB_in and UB_in define the contracted parameter space (contracted in direction p). The outputs are the modified bounds LB_out and UB_out . Points drawn from $\{\theta \in \mathbb{R}^5 : LB_{\text{in}} \leq \theta \leq UB_{\text{in}}\}$ are unlikely to satisfy the polytope constraints. In particular, if

$$
LB_{\text{in}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad UB_{\text{in}} = \begin{bmatrix} 10^{-4} \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

then it is likely that components $3 - 5$ violate the condition $\theta_k \in [0, \min\{\theta_1 \, \theta_2\}]$. To resolve this issue the upper bound is modified, so that $UB_{\text{out},1} = UB_{\text{in},1}$, $UB_{\text{out},2} = UB_{\text{in},2}$, and $UB_{\text{out},k} = \min\{UB_{\text{in},1}, UB_{\text{in},2}, UB_{\text{in},k}\}$ for $k = 3, 4, 5$ (see Lines 39-44 in bound_transform.m). The lower bound is unchanged. The box B defined by LB_{out} and UB_{out} contains the contracted parameter space and retains a good volume ratio. The modifications to bound_transform.m are model specific, and depend on the constraints $A_\theta \theta \leq b_\theta$.

If the parameter space is a polytope, then additional constraints for the linear program that computes $\hat{c}(\cdot)$ are required. These constraints are determined by the user-specified function Rho_Polytope_Box. Recall that we require $\lambda \in \sqrt{n}(\Theta - \theta) \cap \rho B^d$. The constraint $\lambda_k \in [-\rho, \rho]$ is already included in KMS_33_Coverage. For DGP 8, the following constraints need to be added:

$$
\lambda_k \leq \sqrt{n}(1 - \theta_k), k = 1, 2
$$
$$
-\lambda_k \leq \sqrt{n}(0 - \theta_k), k = 1, 2, 3, 4, 5
$$
$$
-\lambda_1 + \lambda_k \leq -\sqrt{n}(-\theta_1 + \theta_k), k = 3, 4, 5
$$
$$
-\lambda_2 + \lambda_k \leq -\sqrt{n}(-\theta_2 + \theta_k), k = 3, 4, 5.
$$

Observe that the constraint $-\lambda_1 + \lambda_k \leq -\sqrt{n}(-\theta_1 + \theta_k)$ is implied by the condition $\theta_k \leq \min\{\theta_1, \theta_2\}$. These $S = 13$ constraints are specified in `Rho_Polytope_Box`. In the `CVXGEN C` code generator, we set $S = 13$ for this DGP.