2.4 November 27–28, 2018

Last commit: `75f77a179`

You had misunderstood the instructions from June. The overall factor multiplying the minimum observational equivalence should never be below 1. In the cases where the model is correctly specified, but there is statistical noise, a factor below 1 makes the LP problem infeasible.

Correctly specified model, $N = 1000$

A small scale simulation with 500 iterations and $N = 1000$ observations is summarized in Table 1 and Figure 1.

Table 1: Summary of simulations, $N = 1000$

τ	Avg. lower	Avg. upper	Lower s.e.	Upper s.e.
0.05	5.78	9.73	2.65	2.16
0.5	4.94	10.45	1.91	1.54
1	4.57	10.68	1.80	1.41
1.5	4.52	10.77	1.80	1.35

Notes: Sample size is 1,000. Each simulation consists of 500 iterations. The model is correctly specified in all simulations.

Correctly specified model, $N = 500$

Now reduce the sample size to $N = 500$. In general, the standard errors are greater. The bounds are also wider on average. The left tail of the distribution of lower bound estimates grows quite a bit, but this is not the case for the right tail of the upper bound estimate. The results are summarized in Table 2 and Figure 2.

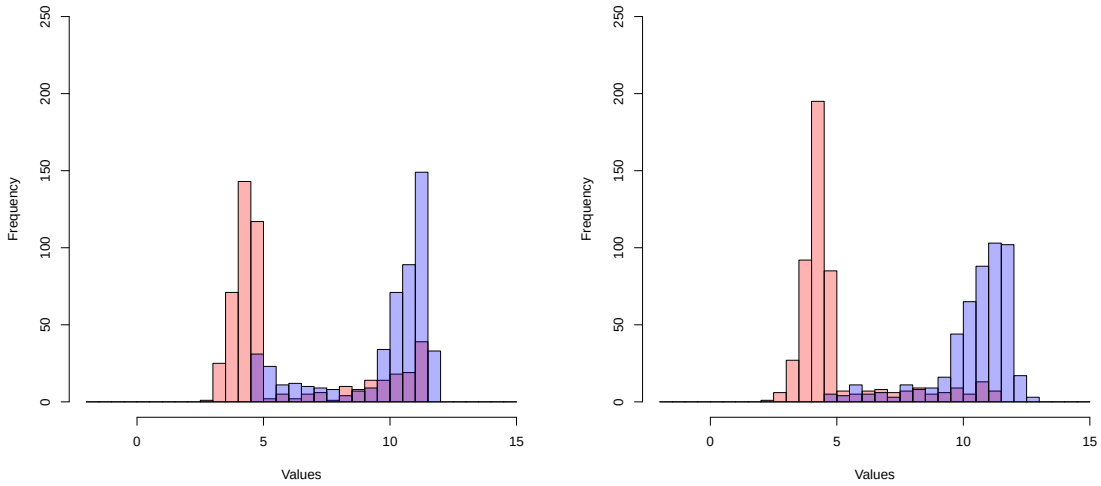
Misspecified model, $N = 1000$

Now you misspecify the model. The correct specification is

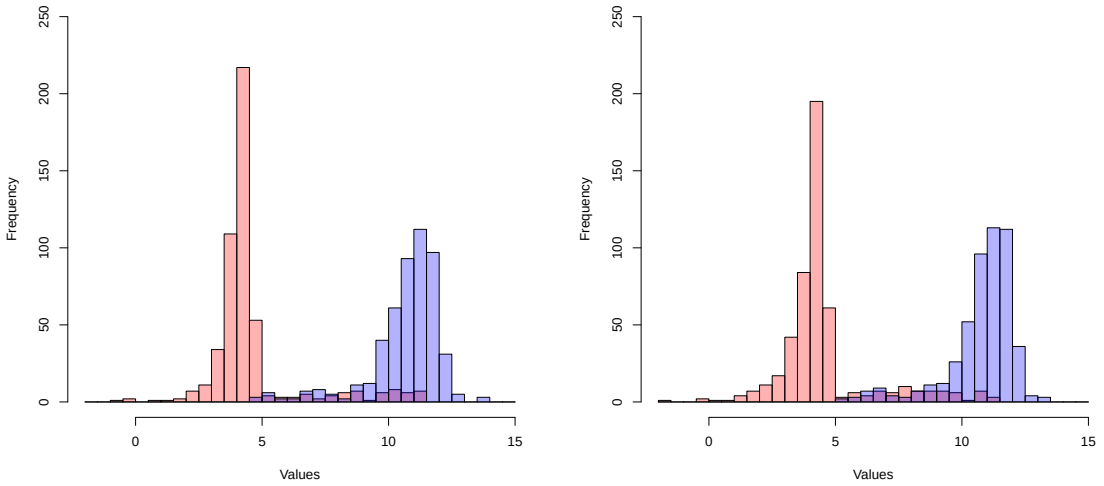
```
ivmte(...
      m0 = ~ 0 + u,
      m1 = ~ 1 + u,
      ...)
```

Figure 1: Distribution of bounds under correct specification, $N = 1000$

(a) $\tau = 0.05$ (b) $\tau = 0.5$



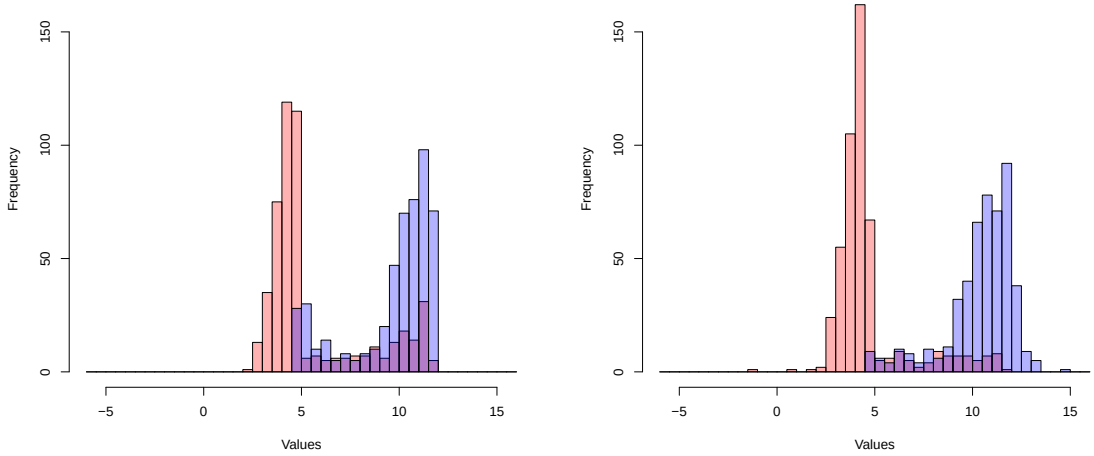
(c) $\tau = 1$ (d) $\tau = 1.5$



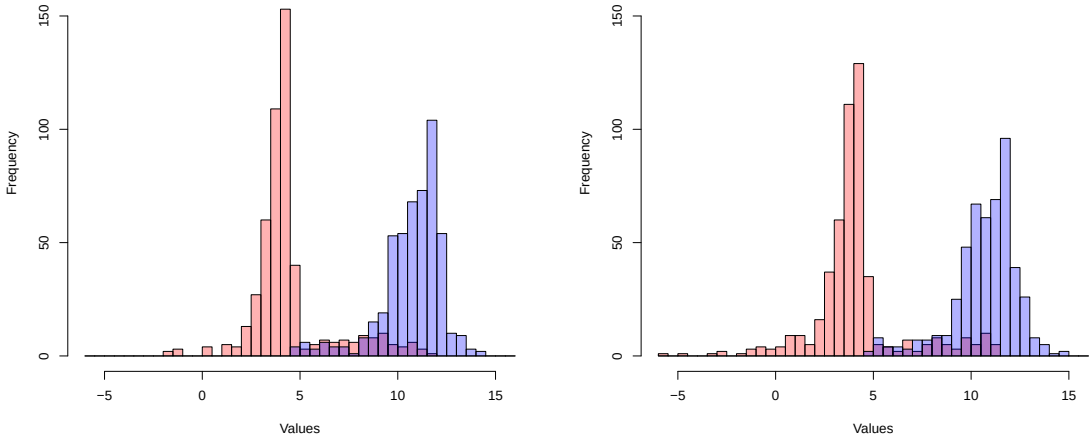
Note: The red bars represent the distribution of the lower bound estimates in the simulation, and the blue bars represent the distribution of the upper bound estimates. Each bar has width 0.5. Sample size is 1,000. Each simulation consists of 500 iterations. The model is incorrectly specified.

Figure 2: Distribution of bounds under correct specification, $N = 500$

(a) $\tau = 0.05$ (b) $\tau = 0.5$



(c) $\tau = 1$ (d) $\tau = 1.5$



Note: The red bars represent the distribution of the lower bound estimates in the simulation, and the blue bars represent the distribution of the upper bound estimates. Each bar has width 0.5. Sample size is 500. Each simulation consists of 500 iterations. The model is incorrectly specified.

Table 2: Summary of simulations, $N = 500$

τ	Avg. lower	Avg. upper	Lower s.e.	Upper s.e.
0.05	5.60	9.67	2.58	2.17
0.5	4.67	10.41	1.93	1.72
1	4.43	10.72	2.00	1.59
1.5	4.11	10.71	2.32	1.65

Notes: Sample size is 500. Each simulation consists of 500 iterations. The model is correctly specified in all simulations.

but you instead declare

```
ivmte(...
      m0 = ~ 0 + u + I(u^2),
      m1 = ~ 1 + u,
      ...)
```

The population bound is $(1.63, 11.36)$, and the minimum deviation from observational equivalence is 0.

Table 3 and Figure 3 summarize these simulations.

Table 3: Summary of simulations, $N = 1000$, misspecified model

τ	Avg. lower	Avg. upper	Lower s.e.	Upper s.e.
0.05	1.85	10.56	1.02	0.96
0.5	1.81	10.58	1.00	0.95
1	1.86	10.52	1.05	1.02
1.5	1.79	10.58	1.01	0.94

Notes: Sample size is 1,000. Each simulation consists of 500 iterations. The model is incorrectly specified. Specifically, m_0 is declared as `0 + u + I(u ^ 2)` instead of `0 + u`.

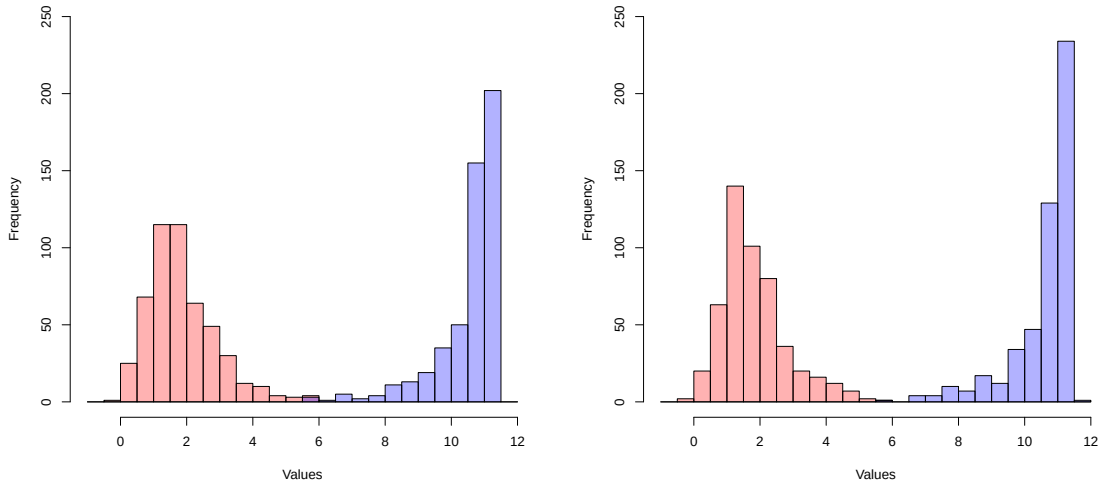
3 `Rcplex` failed estimates

Github issue: `None`

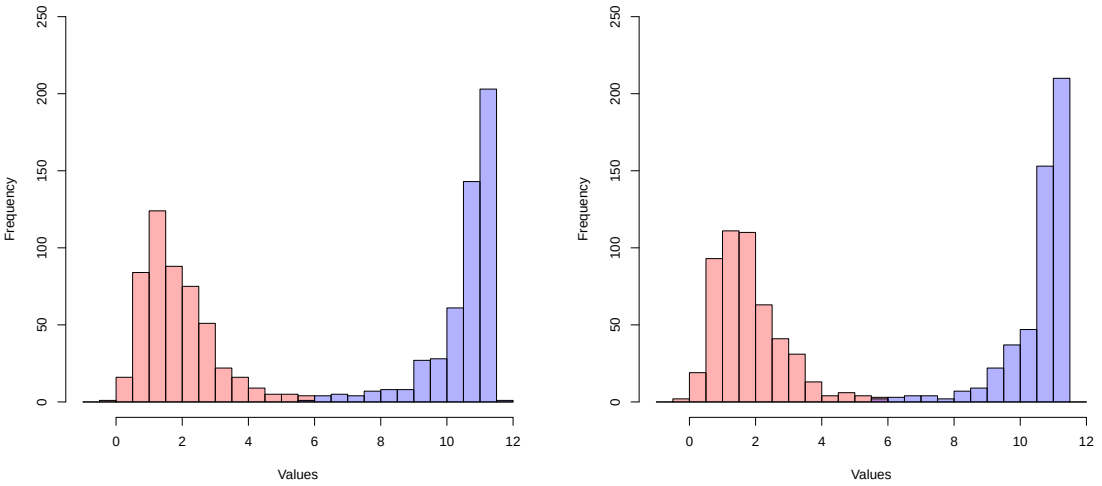
To be clear, the code you are running on the server uses `lpSolve` as the LP solver.

Figure 3: Distribution of bounds under misspecification, $N = 1000$

(a) $\tau = 0.05$ (b) $\tau = 0.5$



(c) $\tau = 1$ (d) $\tau = 1.5$



Note: The red bars represent the distribution of the lower bound estimates in the simulation, and the blue bars represent the distribution of the upper bound estimates. Each bar has width 0.5. Sample size is 1,000. Each simulation consists of 500 iterations. The model is incorrectly specified. Specifically, m_0 is declared as `0 + u + I(u ^ 2)` instead of `0 + u`.