# MEASURING EFFICIENCY IN INFLUENCE DIAGRAM MODELS


Barry R. Cobb

Virginia Military Institute


PGM —2008

# Outline

- Background
- Example—Representation
- Measuring Accuracy, Complexity, and Efficiency
- Description of Experiment
- Example—Solution
- Results
- Summary

# Background

• An *influence diagram* (ID) is a graphical representation of an uncertain decision problem. Ovals are chance nodes, rectangles are decision nodes, and the diamond is the utility node.

• IDs were initially developed as a more compact version of a decision tree.

• The numerical specification of an ID requires conditional probability distributions for chance nodes and a joint utility function.

• A recent paper describes an ID model that accommodates continuous decision variables and non-Gaussian chance variables (Cobb 2007).

# Background

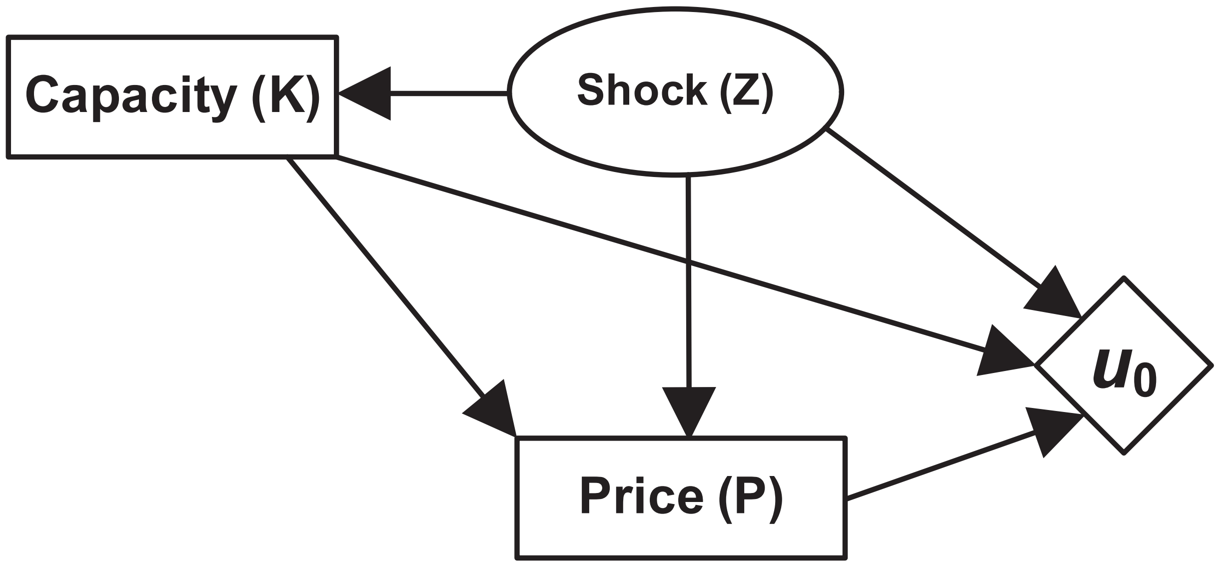
• This paper provides an efficiency measurement that can be used to compare ID models and solution techniques.

• Potential benefits:

↦ 1. Determining whether the additional accuracy gained by using an ID model that accommodates continuous decision variables is worth the incremental computational complexity (as compared to making all decision variables discrete).

↦ 2. Aiding in the design of a decision support system by identifying an ID model that is consistent with a decision maker's desire for more accurate versus less complex (and less costly) solutions.

# Example — Capacity Planning and Pricing under Uncertainty (Göx 2002)




Capacity ($K$) and price ($P$) are decision variables, the random demand “shock” ($Z$) is a chance variable, and $u_0$ is the joint utility function.

# Example—Numerical Details

- Product demand is determined as $Q(p, z) = 12 - p + z$.

- The firm's utility (profit) function is

$$u_0(k, p, z) = \begin{cases} (p-1) \cdot (12 - p + z) - k & \text{if } (12 - p + z) \leq k \\ (p-1) \cdot k - k & \text{if } (12 - p + z) > k \ . \end{cases}$$

- Unit variable cost: $v = 1$; unit capacity cost: $r = 1$.

- $K$ and $P$ are continuous decision variables with state spaces: $\Omega_K = \{k : 0 \leq k \leq 14\}$ and $\Omega_P = \{p : 1 \leq p \leq 9\}$.

- Assume $Z \sim N(0, 1)$

# Example — Analytical Solution

• Firm knows the true value, $Z = z$, of the demand shock $Z$ when it chooses capacity, so it sets $K = 12 - P + z$.

• Göx (2002) finds an analytical solution to the problem with optimal values for $P$ and $K$ of

$$p^* = \Theta_1^*(z) = 2 + \frac{10 + z}{2} \quad \text{and} \quad k^* = \Theta_2^*(z) = \frac{10 + z}{2} \ .$$

• The accuracy of the ID solution can be judged in comparison to the analytical solution.

# Measuring Accuracy

• Mean squared error (MSE) is a measure of the difference between the analytical and ID decision rules.

• $\Theta_2$ is a decision rule for $K = f(Z)$ determined using an ID. The MSE of $\Theta_2$ is

$$MSE = E\left[\left(\Theta_2(z) - \Theta_2^*(z)\right)^2\right] = \int_{\Omega_Z} \phi(z) \cdot (\Theta_2(z) - \Theta_2^*(z))^2 \; dz$$

$\mapsto$ $\phi$ is the density function for $Z$

# Measuring Accuracy

- MSE between the ID decision rule $\Theta_1$ for $P = f(K, Z)$ and the analytical decision rule $\Theta_1^*$ is

$$
\begin{aligned}
MSE = E\left[\left(\Theta_1(\Theta_2(z), z) - \Theta_1^*(z)\right)^2\right] \\
= \int_{\Omega_Z} \phi(z) \cdot (\Theta_1(\Theta_2(z), z) - \Theta_1^*(z))^2 \; dz \; .
\end{aligned}
$$

- The accuracy of an ID model, $\mathcal{A}$, is defined as the sum of the MSEs.

# Measuring Complexity

• ID models in this paper are solved using Mathematica software (`www.wolfram.com`).

• `LeafCount` that gives the total "size" of an expression defined using the `Piecewise` representation, based on applying the `FullForm` function.

• `LeafCount` (denoted by $\mathcal{L}$) will be used to measure complexity by tracking the size of the potentials stored in memory after each combination or marginalization operation (or sub-operation thereof) for the ID solution.

$\mapsto$ `LeafCount` of the potentials affects both the storage required and the subsequent number of calculations.

# Measuring Complexity — Example

- Consider the expression

$$f(z) = \begin{cases} -84.0 + 81.1 \exp\{0.0119(z+3)\} & \text{if } \; -3 \leq z \leq 3\,. \end{cases}$$

- Applying the `Piecewise` and `FullForm` functions in Mathematica to this expression yields

`Piecewise[List[List[ Plus[-84, Times[81.1, Power[E, Times[ 0.0119, Plus[3,` $z$`]]]]]], LessEqual[-3,` $z$`, 3]]], 0] .`

$\mapsto$ Each word, number, or variable in the `FullForm` expression increases the `LeafCount` of the expression by one. In this case, $\mathcal{L}\{f\} = 19$.

# Normalized Measurements

- $\mathcal{A}$ and $\mathcal{C}$ will be normalized onto a common scale to determine the trade-off between accuracy and complexity. Assume $\mathcal{N}_{min} = 1$ and $\mathcal{N}_{max} = 2$.

- The normalized accuracy and complexity measurements are

$$\widehat{\mathcal{A}} = \mathcal{N}_{min} + \frac{(\mathcal{N}_{max} - \mathcal{N}_{min}) \cdot (\mathcal{A} - \underline{\mathcal{A}})}{\overline{\mathcal{A}} - \underline{\mathcal{A}}} \, .$$

$$\widehat{\mathcal{C}} = \mathcal{N}_{min} + \frac{(\mathcal{N}_{max} - \mathcal{N}_{min}) \cdot (\mathcal{C} - \underline{\mathcal{C}})}{\overline{\mathcal{C}} - \underline{\mathcal{C}}} \, .$$

# Measuring Efficiency

• The *efficiency* of the model is calculated as

$$\mathcal{E} = \widehat{\mathcal{A}}^{\alpha} \cdot \widehat{\mathcal{C}}^{1-\alpha}.$$

$\mapsto$ $\alpha$ is assigned by the decision maker to convey an individual preference for solutions that are either more accurate or less complex.

$\mapsto$ If $\alpha > 0.5$, the decision maker values accuracy over complexity, and vice versa.

• If two ID models have equivalent accuracy, the model with a better complexity score will have greater efficiency, and vice versa. Also, there is a diminishing marginal rate of substitution between accuracy and complexity.

## Experiment

• The example will be solved with three types of IDs using a common solution algorithm (the fusion algorithm of Shenoy (1993)):

$\mapsto$ 1. Discrete influence diagram

$\mapsto$ 2. Mixtures of Truncated Exponentials (MTEID) influence diagram (Cobb and Shenoy 2008)

$\mapsto$ 3. Continuous Decision Mixtures of Truncated Exponentials (CDMTEID) influence diagram (Cobb 2007)

• In each method, the number of discrete states used can improve accuracy while increasing complexity. Several alternatives will be considered.

# Example — Representation

• Capacity ($K$) is limited to discrete outcomes in all models. In the CDMTEID model, price ($P$) is maintained as a continuous variables ($K$ is still discrete because continuous decision variables are limited to one continuous parent).

• Use discrete state space $\Omega_K^{(k)} = \{k_1, k_2, \ldots, k_6\}$, i.e. six possible values will be considered for $K$.

• The function $f_1(p) = p$ on the interval $[1, 9]$ is modeled by the MTE potential

$$u_P(p) = -107.056144 + 108.102960 \exp\{0.0089234(p-1)\}$$

# Example — Representation

- The function $f_2(z) = z$ on the interval $[-3, 3]$ is modeled by a similar MTE approximation.

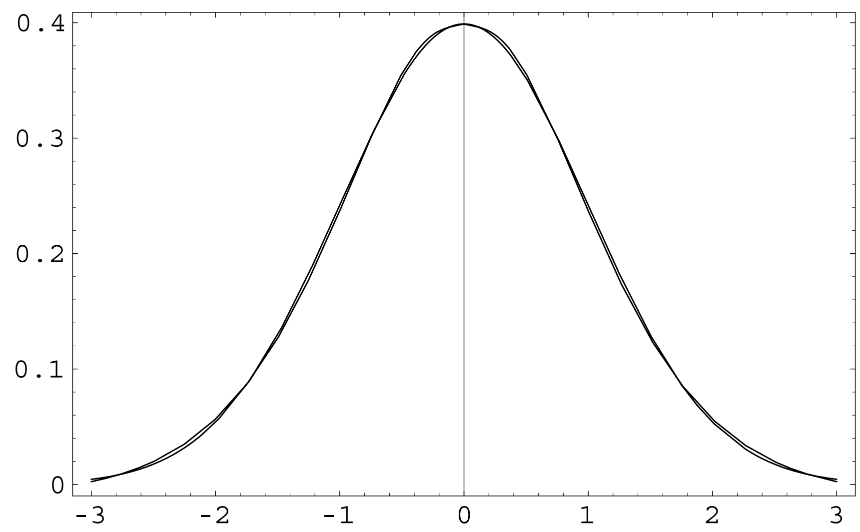
- With $K$ assigned six discrete values, the MTE utility function is defined as

$$u_1(k_t, p, z) = \begin{cases} (u_P(p) - 1) \cdot \\ (12 - u_P(p) + u_Z(z)) - k_t & \text{if } (12 - p + z) \leq k_t \\ (u_P(p) - 1) \cdot k_t - k_t & \text{if } (12 - p + z) > k_t \ , \end{cases}$$

for $t = 1, \ldots, 6$. The result is an MTE potential.

# Example — Representation

- The MTE potential $\phi$ that approximates the distribution (as defined by Cobb and Shenoy (2006)) for the demand shock ($Z$) overlaid on the actual $N(0,1)$ distribution.




- The CDMTEID solution has initial complexity $\mathcal{C}_0 = \mathcal{L}\{\phi\} + \mathcal{L}\{u_1\} = 517$. Other models have different $\mathcal{C}_0$ values.

# Example — CDMTEID Solution

•The elimination sequence employed in the fusion algorithm is $P$, $K$, $Z$.

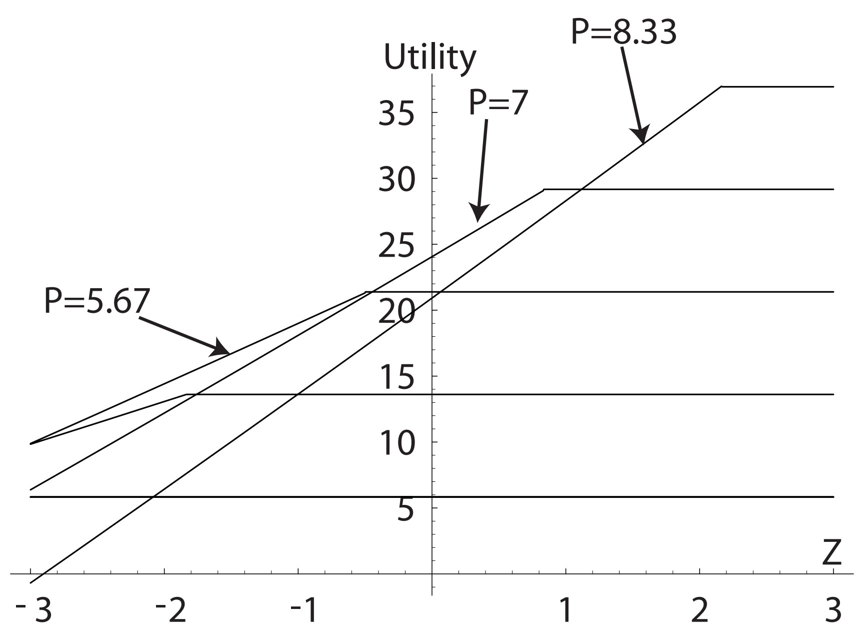
• Price ($P$) is a continuous decision variable; however, the first step in marginalizing this variable is accomplished by using the discrete approximation $\Omega_P^{(d)}$.

$\mapsto$ **Step 1:** The values $p_u$, $u = 1, \ldots, 6$ are used to form the utility functions $u_1(k_t, p_1, z)$, ..., $u_1(k_t, p_6, z)$ for $t = 1, \ldots, 6$. After this step, both these new potentials and the existing MTE utility function $u_1$ remain, so the complexity is

$$\mathcal{C}_1 = \mathcal{C}_0 + \sum_{t=1}^{6} \sum_{u=1}^{6} \mathcal{L}\{u_1(k_t, p_u, z)\} = 1313\,.$$

# Example — CDMTEID Solution

**Step 2:** (for marginalizing $P$) Create a piecewise linear decision rule for $P = f(Z)$ for each value $k_t$, $t = 1, \ldots, 6$. For $K = k_3 = 5.83$, the utility functions $u_1(k_3, p_u, z)$ for $u = 1, \ldots, 6$.



# Example — CDMTEID Solution

**Step 2:** (for marginalizing $P$) We can conclude that $P = 5.67$ is optimal over $[-3, -0.45)$, $P = 7$ is optimal over $[-0.45, 1.15)$, and $P = 8.33$ is optimal over $[1.15, 3]$. These values are used to create the piecewise linear decision rule

$$P(z) = \Theta_{1,3}(z) = \begin{cases} 6.775100 + 0.642570z & \text{if } -3 \leq z < -0.35 \\ 6.729469 + 0.772947z & \text{if } 0.35 \leq z < 2.9375 \\ 9 & \text{if } 2.9375 \leq z \leq 3\,. \end{cases}$$

Similar decision rules $\Theta_{1,1}$, ..., $\Theta_{1,6}$ are determined corresponding to values $k_t$, $t = 1, \ldots, 6$. When combined, these functions form $\Theta_1$ with $\mathcal{L}\{\Theta_1\} = 259$. Since $\phi$ and $u_1$ also remain, $\mathcal{C}_2 = 517 + 259 = 776$.

# Example — CDMTEID Solution

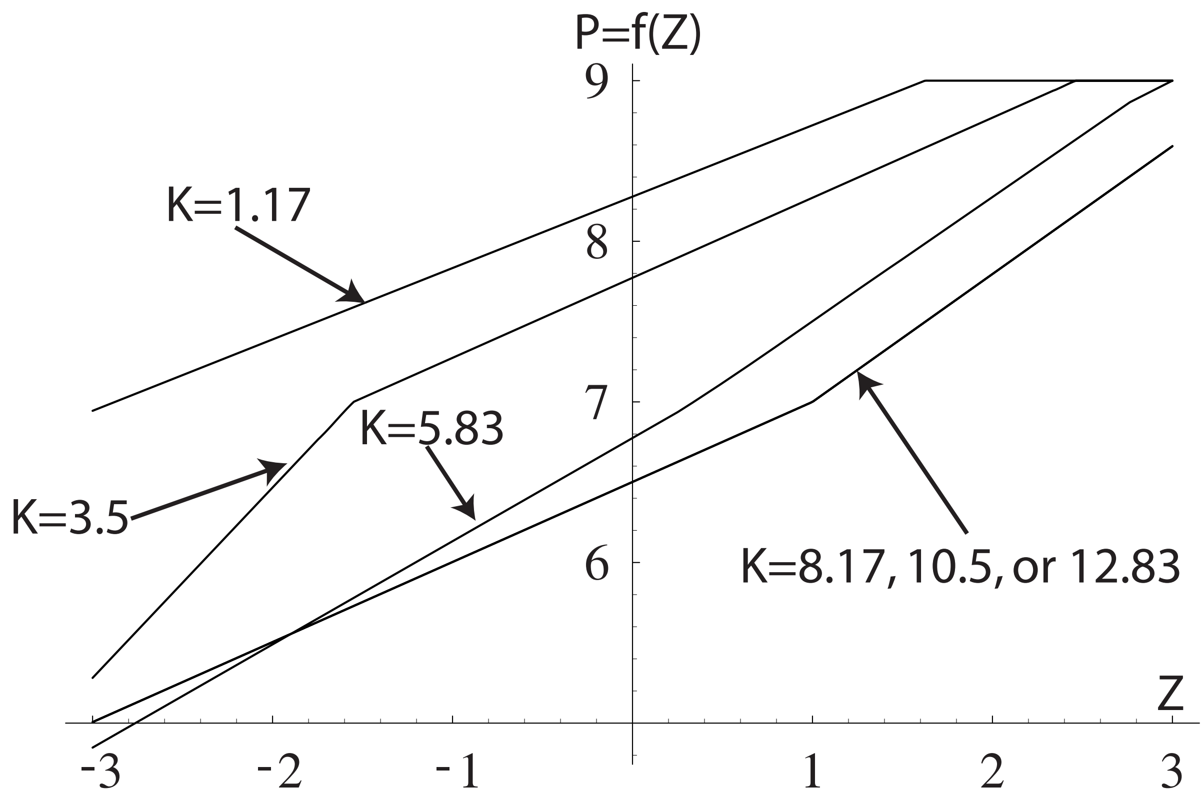
**Step 3:** (for marginalizing $P$) Substitute $\Theta_1$ into $u_1$ to form $u_2(k_t, z) = u_1(k_t, \Theta_{1,t}(z), z)$, for $t = 1, \ldots, 6$. With $\phi$ and $u_2$ as the remaining potentials in the network, the complexity stands at

$$\mathcal{C}_3 = \mathcal{L}\{\phi\} + \sum_{t=1}^{6} \mathcal{L}\{u_2(k_t, z)\} = 61 + 530 = 591$$

• The variables remaining in the model are $K$ (a discrete decision variable) and $Z$ (a continuous chance variable).

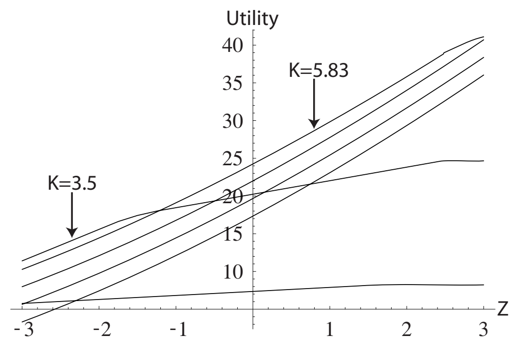
# Example — CDMTEID Solution

- The decision rule $\Theta_1$ for $P = f(K, Z)$.



# Example — CDMTEID Solution

**Marginalizing $K$ (Step 1):** A plot of $u_2(k_1, z)$, ..., $u_2(k_6, z)$ shows that $u_2(3.5, z) \approx u_2(5.83, z)$ at $Z = -1.25$.




The resulting decision rule $\Theta_2$ specifies that $K = 3.5$ if $-3 \leq z < -1.25$ and $K = 5.83$ if $-1.25 \leq z \leq 3$. After creating this decision rule, the complexity of the model is

$$\mathcal{C}_4 = \mathcal{L}\{\phi\} + \mathcal{L}\{u_2\} + \mathcal{L}\{\Theta_2\} = 610$$

# Example — CDMTEID Solution

**Marginalizing $K$ (Step 2):** Calculate $u_3(z) = u_2(\Theta_2(z), z)$. The resulting complexity is $\mathcal{C}_5 = 278$.

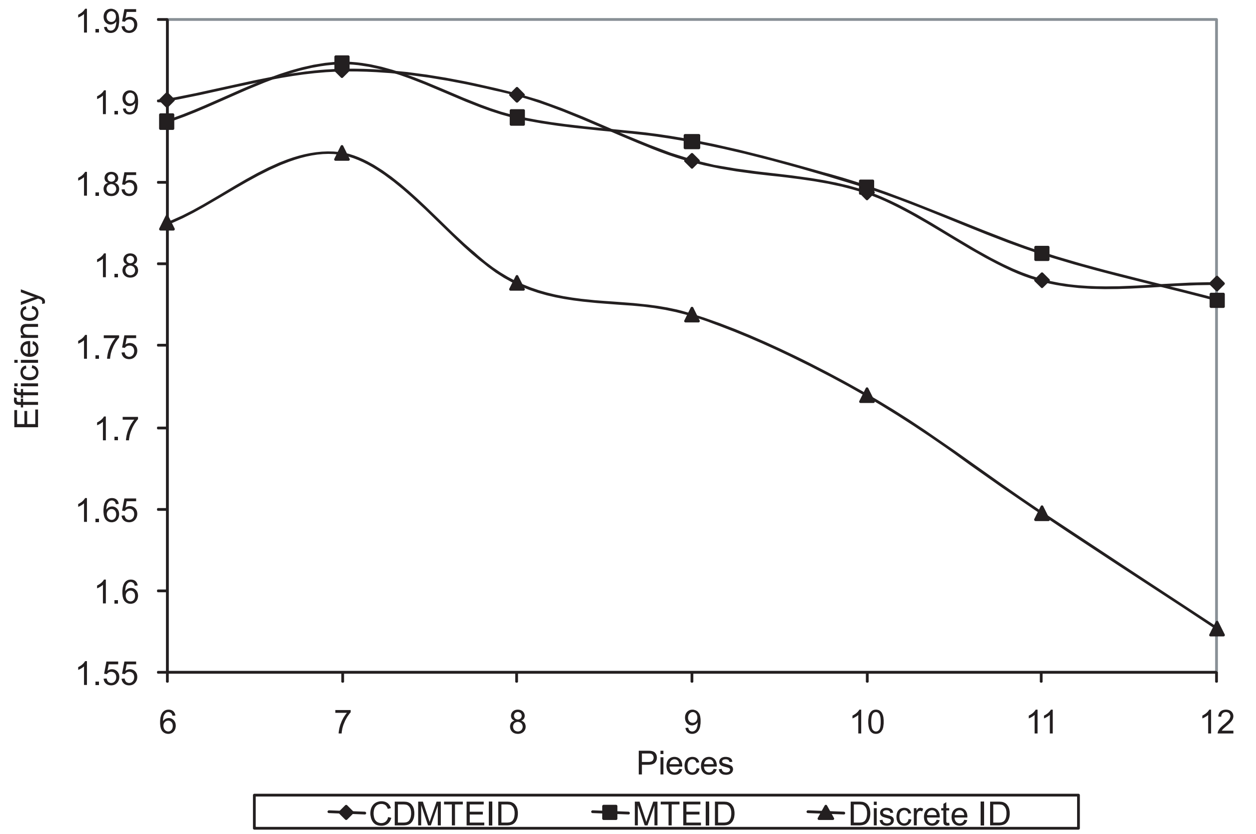
• To remove $Z$, combine $\phi$ and $u_3$, with the resulting complexity $\mathcal{C}_6 = \mathcal{L}\{(\phi \otimes u_3)\} = 569$. Integrating the result over the state space of $Z$ completes the solution. The total complexity of the ID model is

$$\mathcal{C} = \sum_{i=0}^{6} \mathcal{C}_i = 4654\ .$$

• The MSE of the CDMTEID solution is calculated as $\mathcal{A} = 0.7760$

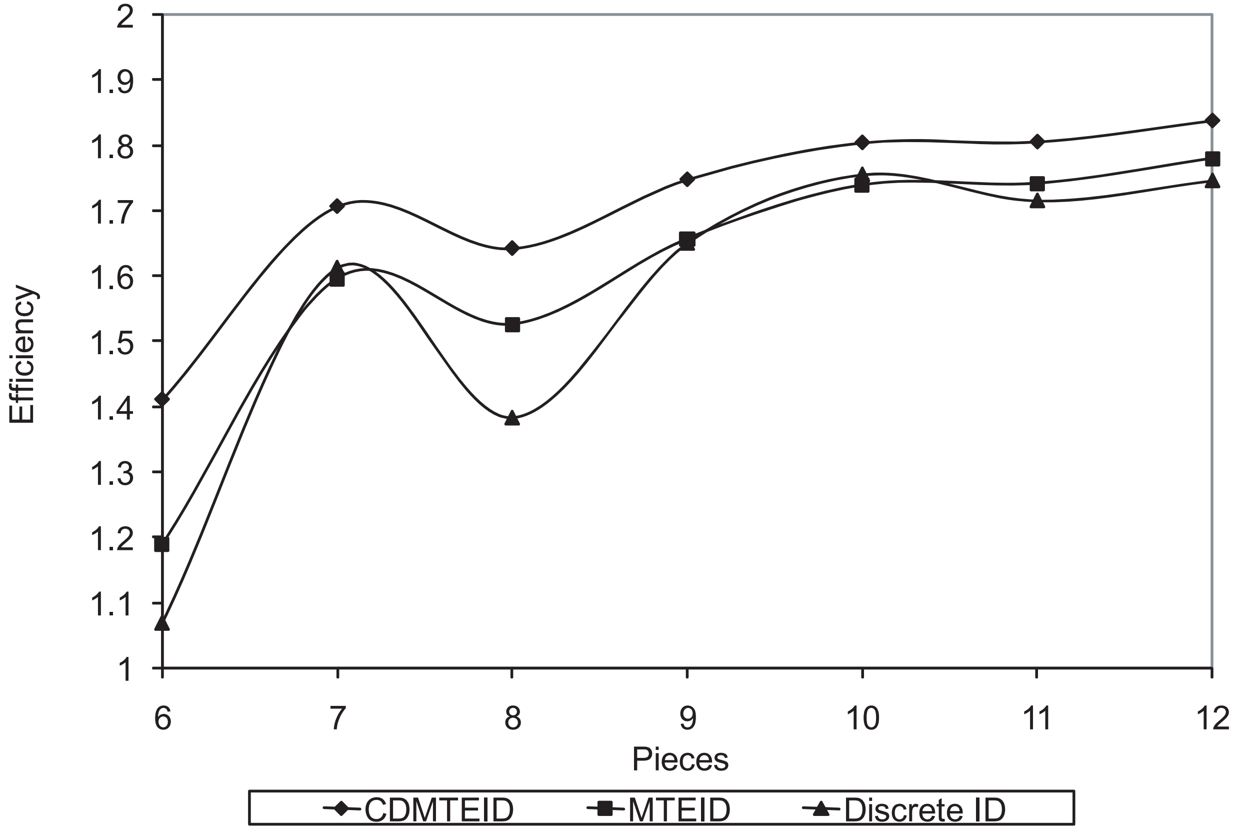
# Results

- Efficiency scores with $\alpha = 0.1$.

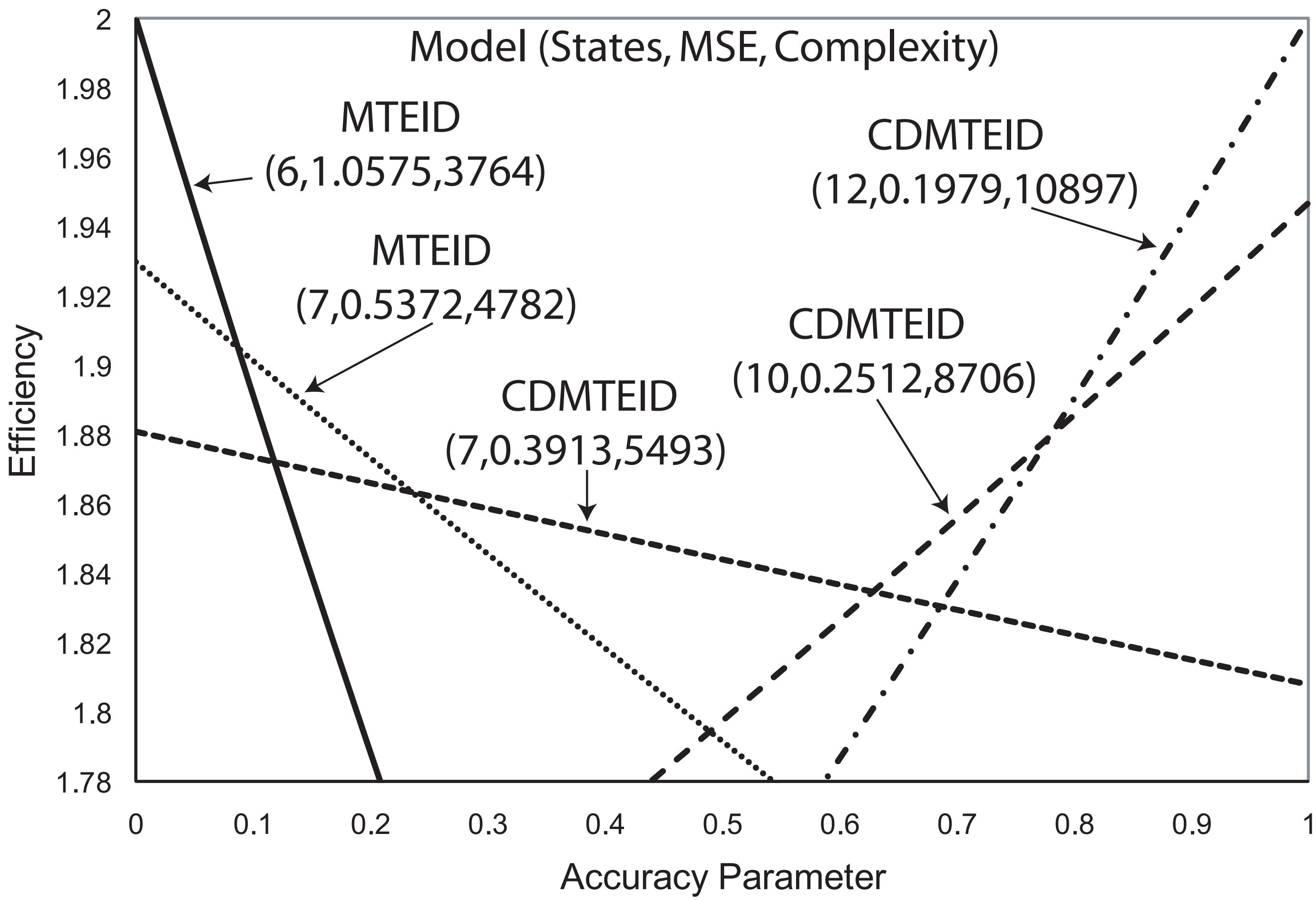


# Results

- Efficiency scores with $\alpha = 0.9$.



# Results



Model (States, MSE, Complexity)
MTEID
(6,1.0575,3764)
MTEID
(7,0.5372,4782)
CDMTEID
(7,0.3913,5493)
CDMTEID
(10,0.2512,8706)
CDMTEID
(12,0.1979,10897)
Efficiency
Accuracy Parameter
2
1.98
1.96
1.94
1.92
1.9
1.88
1.86
1.84
1.82
1.8
1.78
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

# Summary

• When accuracy is a low priority (e.g., $\alpha = 0.1$), efficiency decreases with additional discrete states.

$\mapsto$ MTEID and CDMTEID provide similar efficiency.

• When accuracy is a low priority (e.g., $\alpha = 0.9$), efficiency increases with additional discrete states.

$\mapsto$ CDMTEID provides the best model.

• Using an ID that creates continuous decision rules can be worth the additional complexity, subject to decision maker preferences.