Information Factor: A One Factor Benchmark Model for Asset Pricing

Anisha Ghosh[◇] Christian Julliard[*] Alex Taylor[†]

[◇]Carnegie Mellon University

[*]London School of Economics and CEPR

[†]University of Manchester

ICEF-LFE Workshop in Financial Economics – November 7th 2014

The Bigger Picture

- Asset prices incorporate information about the SDF and the physical probability distribution (e.g. Ross (2013)) i.e. $\mathbb{Q}$

But: asset prices alone are not sufficient to identify (without additional restrictions) the SDF (e.g. Borovička, Hansen and Scheinkman (2014))

⇒ Two large literatures with little intersection:
1. Recovering $\mathbb{Q}$ (mostly from options, or term structure).
2. Identifying sources of (empirically) priced risk – aka risk factors.

We bridge the two and show that we can jointly estimate from the data the in-sample $\mathbb{Q}$ and the sources of priced risk, and project the SDF and priced risk out-of-sample for pricing and investment purposes.

The Bigger Picture

- Asset prices incorporate information about the SDF and the physical probability distribution (e.g. Ross (2013)) i.e. $\mathbb{Q}$

But: asset prices alone are not sufficient to identify (without additional restrictions) the SDF (e.g. Borovička, Hansen and Scheinkman (2014))

$\Rightarrow$ Two large literatures with little intersection:

1. Recovering $\mathbb{Q}$ (mostly from options, or term structure).
2. Identifying sources of (empirically) priced risk – aka risk factors.

We bridge the two and show that we can jointly estimate from the data the in-sample $\mathbb{Q}$ and the sources of priced risk, and project the SDF and priced risk out-of-sample for pricing and investment purposes.

The Bigger Picture

- Asset prices incorporate information about the SDF and the physical probability distribution (e.g. Ross (2013)) i.e. $\mathbb{Q}$

But: asset prices alone are not sufficient to identify (without additional restrictions) the SDF (e.g. Borovička, Hansen and Scheinkman (2014))

⇒ Two large literatures with little intersection:

1. Recovering $\mathbb{Q}$ (mostly from options, or term structure).
2. Identifying sources of (empirically) priced risk – aka risk factors.

We bridge the two and show that we can jointly estimate from the data the in-sample $\mathbb{Q}$ and the sources of priced risk, and project the SDF and priced risk out-of-sample for pricing and investment purposes.

The Bigger Picture

- Asset prices incorporate information about the SDF and the physical probability distribution (e.g. Ross (2013)) i.e. $\mathbb{Q}$

But: asset prices alone are not sufficient to identify (without additional restrictions) the SDF (e.g. Borovička, Hansen and Scheinkman (2014))

⇒ Two large literatures with little intersection:

1. Recovering $\mathbb{Q}$ (mostly from options, or term structure).
2. Identifying sources of (empirically) priced risk – aka risk factors.

We bridge the two and show that we can jointly estimate from the data the in-sample $\mathbb{Q}$ and the sources of priced risk, and project the SDF and priced risk out-of-sample for pricing and investment purposes.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

This Paper

1. extract, in a nonparametric fashion (via entropy), the most likely risk neutral probability measure ($\mathbb{Q}$).
2. Use the MLE of $\mathbb{Q}$ to estimate (via an EE) the SDF and projects it out-of-sample: the Information-SDF
3. Construct an out-of-sample mimicking portfolio: Information-Portfolio

Main Findings:

1. I-SDF and I-P deliver better cross-sectional pricing than multi-factor models (i.e. better encoding of pricing anomalies).
2. I-P has high Sharpe ratio, that outperforms standard benchmarks (market, 1/N, Value, Momentum) out-of-sample.
3. Information factors capture novel information: portable α of $3.5\% - 23.8\%$ (with high hedged Sharpe Ratio)
4. results hold for a wide cross section of assets consisting of size, book-to-market-equity, momentum, industry, and long term reversal sorted portfolios.

(Other) Closely Related Literature

- **Recovery of $\mathbb{Q}$, the SDF, and its characteristics**: e.g. Jackwerth and Rubinstein (1996), Ait-Sahalia Lo (1998, 2000), Rosenberg and Engle (2002), Chernov (2003), Ghosh, Julliard and Taylor (2013), Hansen (2014) ....
- **Entropy based inference**: Stutzer (1996), Kitamura and Stutzer (2002), Julliard and Ghosh (2012), GEL literature ...
- **Cross-sectional asset pricing:** Lewellen, Nagel, Shanken (2010), Harvey, Liu, Zhu (2014), Bryzgalova (2014) ...
- **Portfolio Investment:** Markowitz (1952), MacKinlay and Pastor (2000), Goldfard and Iyengar (2003), Jagannathan and Ma (2003), DeMiguel, Uppal, Wang (2007), DeMiguel, Garlappi, Uppal (2007) ...

(Other) Closely Related Literature

- **Recovery of $\mathbb{Q}$, the SDF, and its characteristics**: e.g. Jackwerth and Rubinstein (1996), Ait-Sahalia Lo (1998, 2000), Rosenberg and Engle (2002), Chernov (2003), Ghosh, Julliard and Taylor (2013), Hansen (2014) ....
- **Entropy based inference**: Stutzer (1996), Kitamura and Stutzer (2002), Julliard and Ghosh (2012), GEL literature ...
- **Cross-sectional asset pricing:** Lewellen, Nagel, Shanken (2010), Harvey, Liu, Zhu (2014), Bryzgalova (2014) ...
- **Portfolio Investment:** Markowitz (1952), MacKinlay and Pastor (2000), Goldfard and Iyengar (2003), Jagannathan and Ma (2003), DeMiguel, Uppal, Wang (2007), DeMiguel, Garlappi, Uppal (2007) ...

(Other) Closely Related Literature

- **Recovery of $\mathbb{Q}$, the SDF, and its characteristics**: e.g. Jackwerth and Rubinstein (1996), Ait-Sahalia Lo (1998, 2000), Rosenberg and Engle (2002), Chernov (2003), Ghosh, Julliard and Taylor (2013), Hansen (2014) ....
- **Entropy based inference**: Stutzer (1996), Kitamura and Stutzer (2002), Julliard and Ghosh (2012), GEL literature ...
- **Cross-sectional asset pricing:** Lewellen, Nagel, Shanken (2010), Harvey, Liu, Zhu (2014), Bryzgalova (2014) ...
- **Portfolio Investment:** Markowitz (1952), MacKinlay and Pastor (2000), Goldfard and Iyengar (2003), Jagannathan and Ma (2003), DeMiguel, Uppal, Wang (2007), DeMiguel, Garlappi, Uppal (2007) ...

(Other) Closely Related Literature

- **Recovery of $\mathbb{Q}$, the SDF, and its characteristics**: e.g. Jackwerth and Rubinstein (1996), Ait-Sahalia Lo (1998, 2000), Rosenberg and Engle (2002), Chernov (2003), Ghosh, Julliard and Taylor (2013), Hansen (2014) ....
- **Entropy based inference**: Stutzer (1996), Kitamura and Stutzer (2002), Julliard and Ghosh (2012), GEL literature ...
- **Cross-sectional asset pricing:** Lewellen, Nagel, Shanken (2010), Harvey, Liu, Zhu (2014), Bryzgalova (2014) ...
- **Portfolio Investment:** Markowitz (1952), MacKinlay and Pastor (2000), Goldfard and Iyengar (2003), Jagannathan and Ma (2003), DeMiguel, Uppal, Wang (2007), DeMiguel, Garlappi, Uppal (2007) ...

Outline

1. The Big Picture
   - Introduction
   - (Other) Closely Related Literature
2. Methodology: An Information Theoretic Approach
   - Change of Measure and Relative Entropy
   - Why Entropy?
3. Empirical Analysis
   - Cross-Sectional Pricing
   - The Information Portfolio
   - Portable α: an "Information Anomaly"
4. Conclusion

Appendix

Outline

Appendix

Change of Measures

- Consider the vector of Euler equations (i.e. no arbitrage restrictions)

$$\mathbf{0} = \mathbb{E}\left[M_t \mathbf{R}_t^e\right] \equiv \int M_t \mathbf{R}_t^e d\mathbb{P}$$

where M_t is the SDF, $\mathbf{R}_t^e$ is a vector of excess returns and $\mathbb{P}$ is the physical probability measure, and $\mathbf{0}$ is a vector of zeros.

- Under very weak regularity conditions, we have

$$\mathbf{0} = \int \frac{M_t}{\bar{M}} \mathbf{R}_t^e \, d\mathbb{P} = \int \mathbf{R}_t^e \, d\mathbb{Q} = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{R}_t^e\right]$$

where $\bar{x} := \mathbb{E}\left[x_t\right]$, $\mathbb{Q}$ is the risk neutral measure and $\frac{M_t}{\bar{M}} = \frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon-Nikodym derivative.

Change of Measures

- Consider the vector of Euler equations (i.e. no arbitrage restrictions)

$$\mathbf{0} = \mathbb{E}\left[M_t \mathbf{R}_t^e\right] \equiv \int M_t \mathbf{R}_t^e d\mathbb{P}$$

  where M_t is the SDF, $\mathbf{R}_t^e$ is a vector of excess returns and $\mathbb{P}$ is the physical probability measure, and $\mathbf{0}$ is a vector of zeros.

- Under very weak regularity conditions, we have

$$\mathbf{0} = \int \frac{M_t}{\bar{M}} \mathbf{R}_t^e \, d\mathbb{P} = \int \mathbf{R}_t^e \, d\mathbb{Q} = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{R}_t^e\right]$$

  where $\bar{x} := \mathbb{E}\left[x_t\right]$, $\mathbb{Q}$ is the risk neutral measure and $\frac{M_t}{\bar{M}} = \frac{d\mathbb{Q}}{d\mathbb{P}}$ is the Radon-Nikodym derivative.

- Given a set of excess returns data, we can estimate $\mathbb{Q}$ as

$$\hat{\mathbb{Q}} = \arg\min_{\mathbb{Q}} D\left(\mathbb{P}||\mathbb{Q}\right) \equiv \arg\min_{\mathbb{Q}} \int \ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P} \text{ s.t. } \mathbf{0} = \int \mathbf{R}_t^e d\mathbb{Q}.$$

$\Rightarrow$ The above is a relative entropy (or KLIC) minimization, under the asset pricing restrictions for the cross section of returns.

Note: $D\left(\mathbb{P}||\mathbb{Q}\right) \geq 0$ and it is measured in bits of information

Also: Since relative entropy is not symmetric, we can also use $D\left(\mathbb{Q}||\mathbb{P}\right)$.

- Given a set of excess returns data, we can estimate $\mathbb{Q}$ as

$$\hat{\mathbb{Q}} = \arg\min_{\mathbb{Q}} D\left(\mathbb{P}||\mathbb{Q}\right) \equiv \arg\min_{\mathbb{Q}} \int \ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P} \text{ s.t. } \mathbf{0} = \int \mathbf{R}_t^e d\mathbb{Q}.$$

$\Rightarrow$ The above is a relative entropy (or KLIC) minimization, under the asset pricing restrictions for the cross section of returns.

Note: $D\left(\mathbb{P}||\mathbb{Q}\right) \geq 0$ and it is measured in bits of information

Also: Since relative entropy is not symmetric, we can also use $D\left(\mathbb{Q}||\mathbb{P}\right)$.

- Given a set of excess returns data, we can estimate $\mathbb{Q}$ as

$$\hat{\mathbb{Q}} = \arg\min_{\mathbb{Q}} D\left(\mathbb{P}||\mathbb{Q}\right) \equiv \arg\min_{\mathbb{Q}} \int \ln\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P} \text{ s.t. } \mathbf{0} = \int \mathbf{R}_t^e d\mathbb{Q}.$$

$\Rightarrow$ The above is a relative entropy (or KLIC) minimization, under the asset pricing restrictions for the cross section of returns.

Note: $D\left(\mathbb{P}||\mathbb{Q}\right) \geq 0$ and it is measured in bits of information

Also: Since relative entropy is not symmetric, we can also use $D\left(\mathbb{Q}||\mathbb{P}\right)$.

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}^e_t \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \equiv \mathbb{E}[\ln p(\mathbf{z})] - \mathbb{E}[\ln q(\mathbf{z})]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\underset{q}{\arg\max}\, \mathbb{E}[\ln q(\mathbf{z})]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}$.

  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.

- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T\ \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}^e_t \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \equiv \mathbb{E}\left[\ln p(\mathbf{z})\right] - \mathbb{E}\left[\ln q(\mathbf{z})\right]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\arg\max_q \mathbb{E}\left[\ln q(\mathbf{z})\right]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}$.

  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.

- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T \; \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}^e_t \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \equiv \mathbb{E}[\ln p(\mathbf{z})] - \mathbb{E}[\ln q(\mathbf{z})]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\arg\max_q \mathbb{E}[\ln q(\mathbf{z})]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}$.

  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.

- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T \ \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}^e_t \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \equiv \mathbb{E}[\ln p(\mathbf{z})] - \mathbb{E}[\ln q(\mathbf{z})]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\arg\max_q \mathbb{E}[\ln q(\mathbf{z})]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}$.
  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.

- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T \ \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}^e_t \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \equiv \mathbb{E}\left[\ln p(\mathbf{z})\right] - \mathbb{E}\left[\ln q(\mathbf{z})\right]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\arg\max_{q} \mathbb{E}\left[\ln q(\mathbf{z})\right]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}$.
  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.
- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T \ \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy?

I: Maximum Likelihood (i.e. not only an Extremum Estimator):

- Let $\mathbf{z}_t$ be a sufficient statistic for the time t state of the economy, then

$$M : \mathbf{z} \to \mathbb{R}_+ \text{ and } \mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N \Rightarrow M_t \equiv M(\mathbf{z}_t) \text{ and } \mathbf{R}_t^e \equiv \mathbf{R}^e(\mathbf{z}_t)$$

- Moreover, defining with p and q the pdf's of $\mathbb{P}$ and $\mathbb{Q}$

$$D(\mathbb{P}||\mathbb{Q}) \equiv \int \ln\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z}) d\mathbf{z} \equiv \mathbb{E}\left[\ln p(\mathbf{z})\right] - \mathbb{E}\left[\ln q(\mathbf{z})\right]$$

- Hence, $\hat{\mathbb{Q}}$ solves: $\arg\max_{q} \mathbb{E}\left[\ln q(\mathbf{z})\right]$ s.t. $\mathbf{0} = \int \mathbf{R}^e(\mathbf{z}) q(\mathbf{z}) d\mathbf{z}$.
  $\Rightarrow$ non parametric MLE (Owen (2001)) of the R-N measure.

- to recover M_t via the Radon-Nikodym derivative simply note that the MLE of p is $p(\mathbf{z}_t) = 1/T \; \forall t = 1, ..., T$

Note: similar argument for $D(\mathbb{Q}||\mathbb{P})$

Why Minimal Entropy? cont'd

- II: Naturally imposes non negativity of the pricing kernel
- III: Can choose any number of assets and don't need a decomposition of M into short and long run components (cf. Alvarez-Jermann), nor it requires a continuum of option data (cf. Ross)
- IV: *Occam's razor*: $\hat{\mathbb{Q}}$ adds to the physical measure the minimum amount of extra information needed to price assets.
- V: Appropriate for capturing tail risk (Brown and Smith (1986))
- VI: Straightforward to add conditional information (i.e. scale the moment function) and/or orthogonality restrictions.
- VII: Simple to construct confidence bands (both asymptotic and by "clever" bootstrap)

VIII: Numerically simple via **duality** (e.g. Csiszar (1975))

1. If $\hat{\mathbb{Q}} = \arg\min\ D(\mathbb{P}||\mathbb{Q})$, we have that (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}_t^e) = \frac{1}{T(1+\hat{\theta}'\mathbf{R}_t^e)}, \quad \forall t \tag{1}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} - \sum_{t=1}^{T} \log(1+\theta'\mathbf{R}_t^e), \tag{2}$$

2. Similarly, if $\hat{\mathbb{Q}} = \arg\min\ D(\mathbb{Q}||\mathbb{P})$, (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}_t^e) = \frac{e^{\hat{\theta}'\mathbf{R}_t^e}}{\sum_{t=1}^{T} e^{\hat{\theta}'\mathbf{R}_t^e}}, \quad \forall t \tag{3}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{T} \sum_{t=1}^{T} e^{\theta'\mathbf{R}_t^e}, \tag{4}$$

VIII: Numerically simple via **duality** (e.g. Csiszar (1975))

1. If $\hat{\mathbb{Q}} = \arg\min \ D(\mathbb{P}||\mathbb{Q})$, we have that (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}^e_t) = \frac{1}{T(1 + \hat{\theta}'\mathbf{R}^e_t)}, \quad \forall t \tag{1}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} - \sum_{t=1}^{T} \log(1 + \theta'\mathbf{R}^e_t), \tag{2}$$

2. Similarly, if $\hat{\mathbb{Q}} = \arg\min \ D(\mathbb{Q}||\mathbb{P})$, (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}^e_t) = \frac{e^{\hat{\theta}'\mathbf{R}^e_t}}{\sum_{t=1}^{T} e^{\hat{\theta}'\mathbf{R}^e_t}}, \quad \forall t \tag{3}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{T} \sum_{t=1}^{T} e^{\theta'\mathbf{R}^e_t}, \tag{4}$$

VIII: Numerically simple via **duality** (e.g. Csiszar (1975))

1. If $\hat{\mathbb{Q}} = \arg\min \; D\left(\mathbb{P}||\mathbb{Q}\right)$, we have that (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}_t^e) = \frac{1}{T(1 + \hat{\theta}'\mathbf{R}_t^e)}, \quad \forall t \tag{1}$$

   where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} - \sum_{t=1}^{T} \log(1 + \theta'\mathbf{R}_t^e), \tag{2}$$

2. Similarly, if $\hat{\mathbb{Q}} = \arg\min \; D\left(\mathbb{Q}||\mathbb{P}\right)$, (up to a scale)

$$\hat{M}_t = M(\hat{\theta}, \mathbf{R}_t^e) = \frac{e^{\hat{\theta}'\mathbf{R}_t^e}}{\sum_{t=1}^{T} e^{\hat{\theta}'\mathbf{R}_t^e}}, \quad \forall t \tag{3}$$

   where $\hat{\theta} \in \mathbb{R}^N$ is the Lagrange multiplier that solves the dual:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{T} \sum_{t=1}^{T} e^{\theta'\mathbf{R}_t^e}, \tag{4}$$

Out-of-Sample Pricing Factors

Out-of-Sample I-SDF

1. Divide data $\mathbf{R}_t^e \in \mathbb{R}^N$, $t = 0, ..., T-1$, into rolling subsamples of length $\bar{T}$, final date T_i, and constant $T_{i+1} - T_i$, $i = 1, 2, ...$
2. Estimate θ_{T_i} over each i-th sample with data sampled at $t = T_i + 1 - \bar{T}, ..., T_i$
3. O-o-S I-SDF: $M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1, ..., T_{i+1}$

Out-of-Sample I-Portfolio

1. $\forall i$ sub-sample set: $\hat{M}_{i,t} = M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1 - \bar{T}, ..., T_i$
2. Project $\hat{M}_{i,t}$ on the space of excess returns to obtain the portfolio weights $\omega_{T_i} \in \mathbb{R}^N$ (normalised to sum to 1)
3. O-o-S I-Portfolio: $R_t^{IP} = \omega_{T_i}' \mathbf{R}_t^e$, $t = T_i + 1, ..., T_{i+1}$

Implementation: $\bar{T}$ = half sample, and $T_{i+1} - T_i = 1$ year (June).

Out-of-Sample Pricing Factors

Out-of-Sample I-SDF

1. Divide data $\mathbf{R}_t^e \in \mathbb{R}^N$, $t = 0, ..., T-1$, into rolling subsamples of length $\bar{T}$, final date T_i, and constant $T_{i+1} - T_i$, $i = 1, 2, ...$
2. Estimate θ_{T_i} over each i-th sample with data sampled at $t = T_i + 1 - \bar{T}, ..., T_i$
3. O-o-S I-SDF: $M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1, ..., T_{i+1}$

Out-of-Sample I-Portfolio

1. $\forall i$ sub-sample set: $\hat{M}_{i,t} = M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1 - \bar{T}, ..., T_i$
2. Project $\hat{M}_{i,t}$ on the space of excess returns to obtain the portfolio weights $\omega_{T_i} \in \mathbb{R}^N$ (normalised to sum to 1)
3. O-o-S I-Portfolio: $R_t^{IP} = \omega'_{T_i} \mathbf{R}_t^e$, $t = T_i + 1, ..., T_{i+1}$

Implementation: $\bar{T}$ = half sample, and $T_{i+1} - T_i = 1$ year (June).

Out-of-Sample Pricing Factors

Out-of-Sample I-SDF

1. Divide data $\mathbf{R}_t^e \in \mathbb{R}^N$, $t = 0, ..., T-1$, into rolling subsamples of length $\bar{T}$, final date T_i, and constant $T_{i+1} - T_i$, $i = 1, 2, ...$
2. Estimate θ_{T_i} over each i-th sample with data sampled at $t = T_i + 1 - \bar{T}, ..., T_i$
3. O-o-S I-SDF: $M(\hat{\theta}_{T_i}, \mathbf{R}_t^e), t = T_i + 1, ..., T_{i+1}$

Out-of-Sample I-Portfolio

1. $\forall i$ sub-sample set: $\hat{M}_{i,t} = M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1 - \bar{T}, ..., T_i$
2. Project $\hat{M}_{i,t}$ on the space of excess returns to obtain the portfolio weights $\omega_{T_i} \in \mathbb{R}^N$ (normalised to sum to 1)
3. O-o-S I-Portfolio: $R_t^{IP} = \omega'_{T_i}\mathbf{R}_t^e$, $t = T_i + 1, ..., T_{i+1}$

Implementation: $\bar{T}$ = half sample, and $T_{i+1} - T_i = 1$ year (June).

Out-of-Sample Pricing Factors

Out-of-Sample I-SDF

1. Divide data $\mathbf{R}_t^e \in \mathbb{R}^N$, $t = 0, ..., T-1$, into rolling subsamples of length $\bar{T}$, final date T_i, and constant $T_{i+1} - T_i$, $i = 1, 2, ...$
2. Estimate θ_{T_i} over each i-th sample with data sampled at $t = T_i + 1 - \bar{T}, ..., T_i$
3. O-o-S I-SDF: $M(\hat{\theta}_{T_i}, \mathbf{R}_t^e), t = T_i + 1, ..., T_{i+1}$

Out-of-Sample I-Portfolio

1. $\forall i$ sub-sample set: $\hat{M}_{i,t} = M(\hat{\theta}_{T_i}, \mathbf{R}_t^e)$, $t = T_i + 1 - \bar{T}, ..., T_i$
2. Project $\hat{M}_{i,t}$ on the space of excess returns to obtain the portfolio weights $\omega_{T_i} \in \mathbb{R}^N$ (normalised to sum to 1)
3. O-o-S I-Portfolio: $R_t^{IP} = \omega'_{T_i}\mathbf{R}_t^e$, $t = T_i + 1, ..., T_{i+1}$

Implementation: $\bar{T}$ = half sample, and $T_{i+1} - T_i = 1$ year (June).

Outline

Cross-Sectional Pricing

Table 1: 25 FF Portfolios, 1929:04-2010:12

$const.$	λ_{IP}	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
				Panel A: Monthly					
0.003		−0.341				67.0	56.6	37.5	0.077
(5.73)		(−7.06)				[57.4,100]	[54.4,100]	(0.182)	[0.0,0.056]
0.003	0.023					68.6	59.6	37.10	0.072
(5.70)	(7.32)					[52.1,100]	[75.0,100]	(0.096)	[0.0,0.064]
0.011			−0.004			3.97	28.8	71.64	0.128
(3.40)			(−1.41)			[−3.82,59.8]	[19.4,40.9]	(0.000)	[0.045,0.145
0.011			−0.006	0.002	0.004	71.3	40.9	51.46	0.096
(2.50)			(−1.53)	(3.86)	(6.87)	[48.9,89.2]	[31.4,82.8]	(0.003)	[0.040,0.100
0.004	0.025		0.0004	0.003	0.004	86.1	62.2	29.25	0.058
(1.20)	(5.26)		(0.132)	(6.75)	(8.48)	[,]	[,]	(0.083)	[,]
				Panel B: Quarterly					
0.028		−5.46				26.8	30.8	41.31	0.332
(11.33)		(−3.13)				[4.38,100]	[16.9,69.5]	(0.469)	[0.0,0.438]
0.002	0.135					83.7	51.6	28.77	0.227
(1.29)	(11.17)					[86.9,100]	[52.2,100]	(0.533)	[0.0,0.164]
0.024			−0.002			−3.92	8.50	80.90	0.431
(2.79)			(−0.308)			[−4.33,25.7]	[4.88,11.9]	(0.000)	[0.14,0.42]
0.028			−0.015	0.007	0.013	74.7	17.7	59.34	0.351
(2.19)			(−1.15)	(4.95)	(7.20)	[0.52,0.94]	[0.018,0.56]	(0.001)	[0.066,0.37]
0.005	0.108		0.010	0.008	0.012	83.4	46.5	31.06	0.217
(0.403)	(3.71)		(0.768)	(6.60)	(7.85)	[,]	[,]	(0.054)	[,]

Cross-Sectional Pricing cont'd

Table 2: 10 Momentum Portfolios, 1929:04-2010:12

const.	λ_{IP}	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	λ_{MOM}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
Panel A: Monthly										
0.004		−0.27					90.2	68.1	12.37	0.024
(9.99)		(−9.16)					[78.9,100]	[56.8,100]	(0.345)	[0.0,0.031]
0.002	0.034						95.4	83.6	6.37	0.012
(7.80)	(13.66)						[92.0,100]	[84.3,100]	(0.657)	[0.0,0.015]
0.014			−0.009				10.9	−2.0	40.15	0.074
(2.04)			(−1.45)				[−12.4,60.1]	[−11.7,16.9]	(0.000)	[0.026,0.117
0.022			−0.013	−0.011	−0.032	0.007	78.9	2.59	8.81	0.044
(1.27)			(−0.75)	(−0.61)	(−1.18)	(6.02)	[32.9,100]	[−73.1,98.4]	(0.399)	[0.0,0.823]
0.003	0.039		0.004	−0.005	−0.015	0.006	97.7	82.1	2.61	0.006
(0.42)	(4.36)		(0.60)	(−0.80)	(−1.60)	(14.64)	[,]	[,]	(0.625)	[,]
Panel B: Quarterly										
0.008		−1.07					87.3	75.2	8.12	0.056
(6.39)		(−7.93)					[67.9,100]	[72.2,100]	(0.529)	[0.0,0.080]
0.006	0.107						95.4	78.6	7.36	0.047
(5.60)	(9.90)						[86.3,100]	[76.9,100]	(0.581)	[0.0,0.048]
0.038			−0.024				14.4	−6.49	39.55	0.226
(2.40)			(−1.59)				[−11.4,81.2]	[−12.5,13.7]	(0.001)	[0.077,0.381
0.061			−0.050	0.038	0.021	0.022	75.4	1.56	9.55	0.131
(1.08)			(−0.83)	(1.03)	(0.63)	(5.46)	[20.9,96.5]	[−73.6,81.2]	(0.187)	[0.0,1.346]
0.036	0.084		−0.024	0.032	0.0005	0.021	98.4	86.7	1.35	0.014
(2.43)	(3.69)		(−1.54)	(3.28)	(0.05)	(20.80)	[,]	[,]	(0.853)	[,]

Cross-Sectional Pricing cont'd

Table 3: 25 Portfolios Formed on Long-Term Reversal and Size, 1929:04-2010:12

const.	λ_{IP}	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
				Panel A: Monthly					
0.006		−0.22				13.5	61.5	25.07	0.047
(7.06)		(−2.18)				[−0.83,100]	[56.4,100]	(0.583)	[0.0,0.036]
0.002	0.024					72.3	68.0	18.35	0.037
(2.13)	(7.98)					[66.4,100]	[71.0,100]	(0.844)	[0.0,0.007]
0.005			0.002			−1.6	10.1	58.14	0.103
(1.38)			(0.78)			[−4.10,44.7]	[1.83,21.7]	(0.002)	[0.028,0.111]
0.002			0.002	0.001	0.007	74.3	26.1	40.37	0.077
(0.68)			(0.93)	(1.62)	(4.96)	[56.7,100]	[12.4,100]	(0.018)	[0.01,0.070]
−0.002	0.022		0.006	0.003	0.004	84.5	66.4	16.69	0.033
(−0.86)	(5.50)		(2.77)	(3.99)	(3.39)	[,]	[,]	(0.673)	[,]
				Panel B: Quarterly					
0.023		−0.33				−3.94	27.2	54.50	0.291
(11.80)		(−0.300)				[,]	[,]	(0.000)	[,]
0.008	0.075					56.8	56.5	23.96	0.167
(3.13)	(5.71)					[39.0,100]	[56.9,100]	(0.735)	[0.0,0.064]
0.008			0.013			8.7	1.33	68.46	0.372
(1.04)			(1.81)			[,]	[,]	(0.000)	[,]
0.006			0.009	0.005	0.020	77.8	11.96	48.86	0.301
(0.651)			(1.03)	(2.59)	(4.76)	[,]	[,]	(0.000)	[,]
0.002	0.070		0.012	0.011	0.007	86.7	53.0	22.61	0.153
(0.329)	(5.28)		(1.80)	(5.13)	(1.47)	[,]	[,]	(0.309)	[,]

Cross-Sectional Pricing cont'd

Table 4: Small, Large, Growth, Value, Winners, Losers, 10 Industry, 1929:04-2010:12

const.	λ_{IP}	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	λ_{MOM}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
Panel A: Monthly										
0.001		−1.12					94.9	88.0	5.39	0.013
(2.49)		(−16.7)					[,]	[,]	(0.980)	[,]
0.003	0.027						83.0	84.2	9.89	0.019
(7.45)	(8.62)						[,]	[,]	(0.770)	[,]
0.007			−0.002				−3.5	−1.17	62.76	0.112
(2.02)			(−0.70)				[,]	[,]	(0.000)	[,]
0.003			0.002	0.002	0.002	0.009	84.8	40.3	27.23	0.052
(1.25)			(1.09)	(2.42)	(1.92)	(8.78)	[,]	[,]	(0.004)	[,]
−0.003	0.035		0.007	0.003	0.0003	0.008	94.7	84.8	5.73	0.012
(−1.58)	(8.10)		(4.45)	(5.44)	(0.51)	(10.65)	[,]	[,]	(0.838)	[,]
Panel B: Quarterly										
0.014		−3.88					75.8	60.9	17.15	0.145
(12.44)		(−6.92)					[,]	[,]	(0.248)	[,]
0.009	0.100						74.0	79.9	10.85	0.077
(5.81)	(6.61)						[,]	[,]	(0.698)	[,]
0.021			−0.005				−5.1	0.73	67.17	0.358
(2.10)			(−0.52)				[,]	[,]	(0.000)	[,]
0.012			0.004	0.006	0.005	0.027	82.9	37.3	27.86	0.178
(1.47)			(0.45)	(2.50)	(2.01)	(7.97)	[,]	[,]	(0.003)	[,]
−0.002	0.107		0.017	0.008	0.002	0.024	89.6	75.8	8.27	0.062
(−0.26)	(4.62)		(2.16)	(3.89)	(1.10)	(8.43)	[,]	[,]	(0.602)	[,]

Cross-Sectional Pricing cont'd

Table 5: 30 Industry Portfolios, 1929:04-2010:12

const.	λ_{IP}	λ_{sdf}	λ_{Rm}	λ_{SMB}	λ_{HML}	$\overline{R}^2_{OLS}$	$\overline{R}^2_{GLS}$	T^2	q
				Panel A: Monthly					
0.002		−0.15				33.8	50.1	12.61	0.023
(2.58)		(−3.97)				[,]	[,]	(0.994)	[,]
0.001	0.022					52.5	65.6	8.56	0.016
(1.42)	(5.74)					[,]	[,]	(1.00)	[,]
0.006			−0.0001			−3.52	−2.30	26.69	0.047
(3.55)			(−0.12)			[,]	[,]	(0.535)	[,]
0.006			−0.001	0.0004	−0.001	−8.33	−7.32	26.00	0.046
(2.46)			(−0.36)	(0.26)	(−0.75)	[,]	[,]	(0.463)	[,]
0.000	0.025		0.005	−0.0001	−0.001	60.4	64.6	7.55	0.015
(0.023)	(5.79)		(2.93)	(−0.14)	(−1.07)	[,]	[,]	(1.00)	[,]
				Panel B: Quarterly					
0.011		−1.72				5.55	31.4	23.49	0.133
(2.93)		(−1.64)				[,]	[,]	(0.708)	[,]
0.010	0.043					17.3	39.9	20.66	0.116
(3.48)	(2.66)					[,]	[,]	(0.839)	[,]
0.017			0.001			−3.51	2.68	35.16	0.187
(3.55)			(0.127)			[,]	[,]	(0.165)	[,]
0.023			−0.006	0.003	−0.002	−8.40	−1.64	33.57	0.181
(2.51)			(−0.62)	(0.71)	(−0.43)	[,]	[,]	(0.146)	[,]
0.020	0.042		−0.003	0.006	−0.002	26.6	42.9	16.29	0.098
(2.58)	(2.42)		(−0.350)	(1.81)	(−0.562)	[,]	[,]	(0.906)	[,]

The Information Portfolio

Table 8: Summary Statistics of Information Portfolio & Returns

Assets	Mean	Vol	SR	Skew	Kurt	CEQ	Ret-gain
Panel A: Monthly							
Market - Risk Free	0.004	0.045	0.091	−0.567	5.028	0.003	
Value - Growth	0.004	0.029	0.139	−0.034	5.440	0.004	
Momentum Factor	0.007	0.044	0.164	−1.419	13.65	0.006	
R^{IP} (FF25)	0.021 (0.007)	0.073 (0.051)	0.288 (0.128)	0.384 (−0.575)	5.541 (5.589)	0.018 (0.006)	0.008
R^{IP} (10 Momentum)	0.030 (0.004)	0.127 (0.048)	0.235 (0.085)	−0.352 (−0.326)	8.022 (4.793)	0.022 (0.003)	0.007
R^{IP} (25 L-T Reversal & Size)	0.013 (0.007)	0.064 (0.051)	0.206 (0.137)	−0.212 (−0.444)	5.111 (5.865)	0.011 (0.006)	0.003
R^{IP} (S, B, G, V, W, L, 10 Ind.)	0.027 (0.005)	0.088 (0.046)	0.306 (0.106)	−0.679 (−0.490)	6.180 (4.953)	0.023 (0.004)	0.009
R^{IP} (30 Industry)	0.002 (0.005)	0.083 (0.048)	0.018 (0.112)	0.040 (−0.522)	6.318 (5.708)	−0.001 (0.004)	−0.004

Note: $1/N$ portfolio in parenthesis.

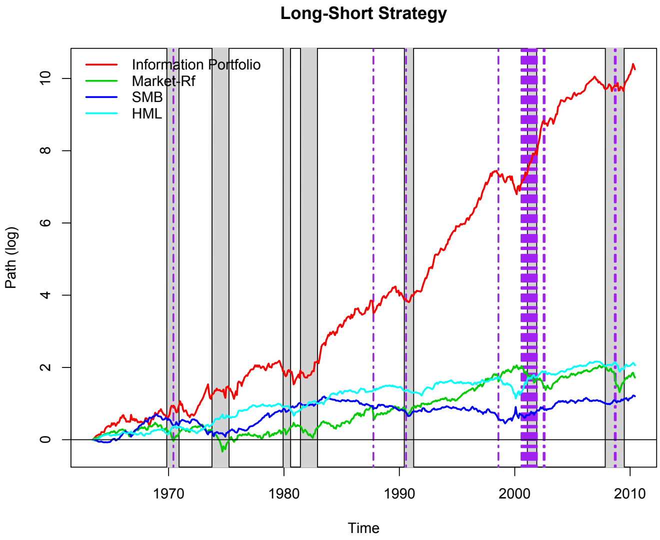
The Information Portfolio

Table 8: Summary Statistics of Information Portfolio & Returns

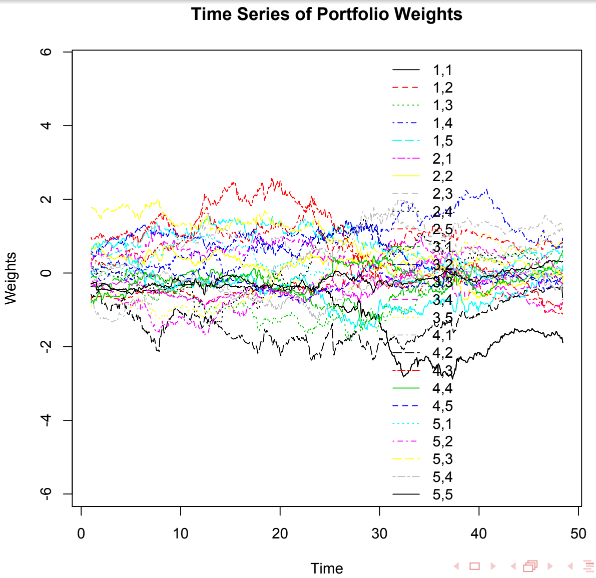
Assets	Mean	Vol	SR	Skew	Kurt	CEQ	Ret-gain
Panel B: Quarterly							
Market - Risk Free	0.013	0.087	0.150	−0.435	3.635	0.009	
Value - Growth	0.012	0.060	0.204	0.109	4.754	0.010	
Momentum Factor	0.020	0.081	0.254	−1.411	10.13	0.017	
R^{IP}	0.080	0.194	0.413	0.410	3.955	0.061	0.022
(FF25)	(0.021)	(0.103)	(0.207)	(−0.183)	(3.576)	(0.016)	
R^{IP}	0.085	0.239	0.354	−0.090	5.295	0.056	0.020
(10 Momentum)	(0.013)	(0.093)	(0.143)	(−0.231)	(3.805)	(0.009)	
R^{IP}	0.042	0.134	0.313	−0.168	3.833	0.033	0.010
(25 L-T Reversal & Size)	(0.023)	(0.104)	(0.220)	(−0.057)	(3.865)	(0.018)	
R^{IP}	0.083	0.173	0.480	0.181	3.463	0.068	0.027
(S, B, G, V, W, L, 10 Ind.)	(0.016)	(0.090)	(0.175)	(−0.315)	(3.794)	(0.012)	
R^{IP}	0.007	0.180	0.041	0.029	2.934	−0.009	−0.013
(30 Industry)	(0.017)	(0.093)	(0.186)	(−0.298)	(3.942)	(0.013)	

Note: $1/N$ portfolio in parenthesis.

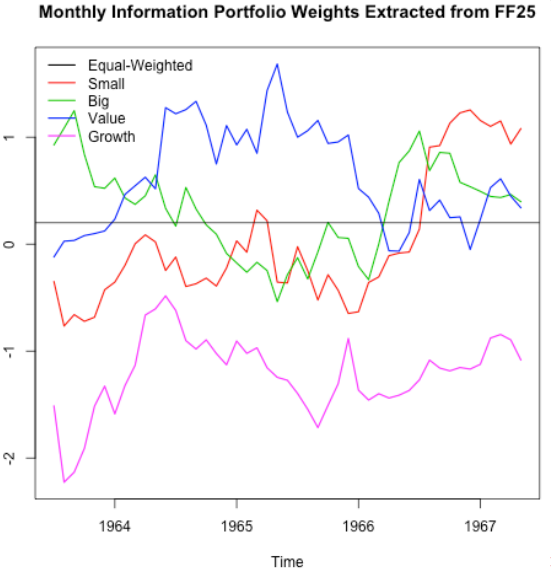
R^{IP} Cumulated Returns



R^{IP} portfolio weights



R^{IP}: S, M , V and G aggregated weights



Portable α: an "Information Anomaly"

Table 7: R^{IP} α With Respect to FF and Momentum Factors, 1929:04-2010:12

Assets	α (%)	β_{Rm}	β_{SMB}	β_{HML}	β_{MOM}	$\overline{R}^2_{OLS}$	Info-Ratio
Panel A: Monthly							
FF25	1.31 (4.89)	0.47 (7.44)	0.21 (2.41)	1.30 (13.82)		26.9	0.211
10 Momentum	1.37 (3.06)	0.94 (9.03)	−0.30 (−2.09)	0.26 (1.67)	1.67 (16.52)	36.1	0.135
25 L-T Reversal & Size	0.69 (2.79)	0.40 (6.92)	0.35 (4.39)	0.92 (10.56)		20.7	0.121
S, B, G, V, W, L, 10 Ind.	0.93 (3.28)	0.59 (9.00)	0.35 (3.90)	1.28 (12.87)	1.25 (19.70)	46.6	0.145
30 Industry	0.29 (0.96)	0.81 (11.43)	−0.87 (−8.79)	−0.59 (−5.57)		28.3	0.041

Annualised α: 3.5%-17.7%

Portable α: an "Information Anomaly" cont'd

Table 7: R^{IP} α With Respect to FF and Momentum Factors, 1929:04-2010:12

Assets	α (%)	β_{Rm}	β_{SMB}	β_{HML}	β_{MOM}	$\overline{R}^2_{OLS}$	Info-Ratio
Panel B: Quarterly							
FF25	5.49 (4.01)	0.59 (3.29)	0.43 (1.72)	1.12 (4.85)		15.4	0.310
10 Momentum	2.63 (1.77)	1.20 (6.52)	−0.21 (−0.82)	0.43 (1.76)	1.92 (10.76)	42.6	0.147
25 L-T Reversal & Size	2.39 (2.57)	0.50 (4.16)	0.01 (0.06)	0.93 (5.94)		18.0	0.199
S, B, G, V, W, L, 10 Ind.	3.21 (2.77)	0.67 (4.70)	0.29 (1.42)	1.16 (6.15)	1.23 (8.89)	33.6	0.230
30 Industry	1.16 (1.00)	1.02 (6.73)	−1.01 (−4.72)	−0.66 (−3.37)		29.1	0.077

Annualised α: 4.7%-23.8%

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Why does it work?

- The maximum Sharpe Ratio portfolio must price the cross-section of assets (since it pins down the Capital Allocation Line).
- Standard methods for forecasting this portfolio use the asset side i.e. tangency of the Efficient Frontier and the Capital Allocation Line

$\Rightarrow$ They don't work empirically due to measurement error on $N + N(N+1)/2$ parameters (mean plus covariance matrix) e.g. DeMiguel, Garlappi, Uppal (2007)

E.g. : with N=25 this method requires 350 parameters!

- Our method uses the SDF side i.e. tangency of the Indifference Curve to the Capital Allocation Line

$\Rightarrow$ need to estimate only N parameters of the dual solution.

Outline

Appendix

Conclusion

The paper provides a novel (ML) method for the extraction of the SDF, and its mimicking portfolio, that:

- Prices assets out-of-sample as well or better than canonical factor models.
- Identifies a large amount of risk not spanned by traditional factors.
- Provides a tradeable portfolio that:
  1. is statistically indistinguishable from the maximum Sharpe Ratio portfolio;
  2. has very high Sharpe Ratio even when hedged w.r.t. traditional factors;
  3. has low turnover, hence low trading costs.