Hey <AI Assistant>, are we in recession now? A review of Business Cycle Turning Point Detection

Eugene Chang, Emmanuelle Guidetti, Annabelle Mourougane

EC-OECD Workshop on Business and Consumer Surveys
17 October 2024

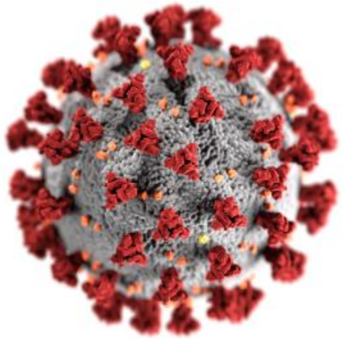
OECD
BETTER POLICIES FOR BETTER LIVES

At a glance

- Economic statistics is evolving due to new challenges and opportunities
- Impacts on traditional early warning indicators of phase changes (*turning points*) between expansion and recession
- Review methods for turning point detection
  - Frameworks for assessment
  - Explore non-mainstream methods and alternative data

Why review turning point methods?

Global Challenges



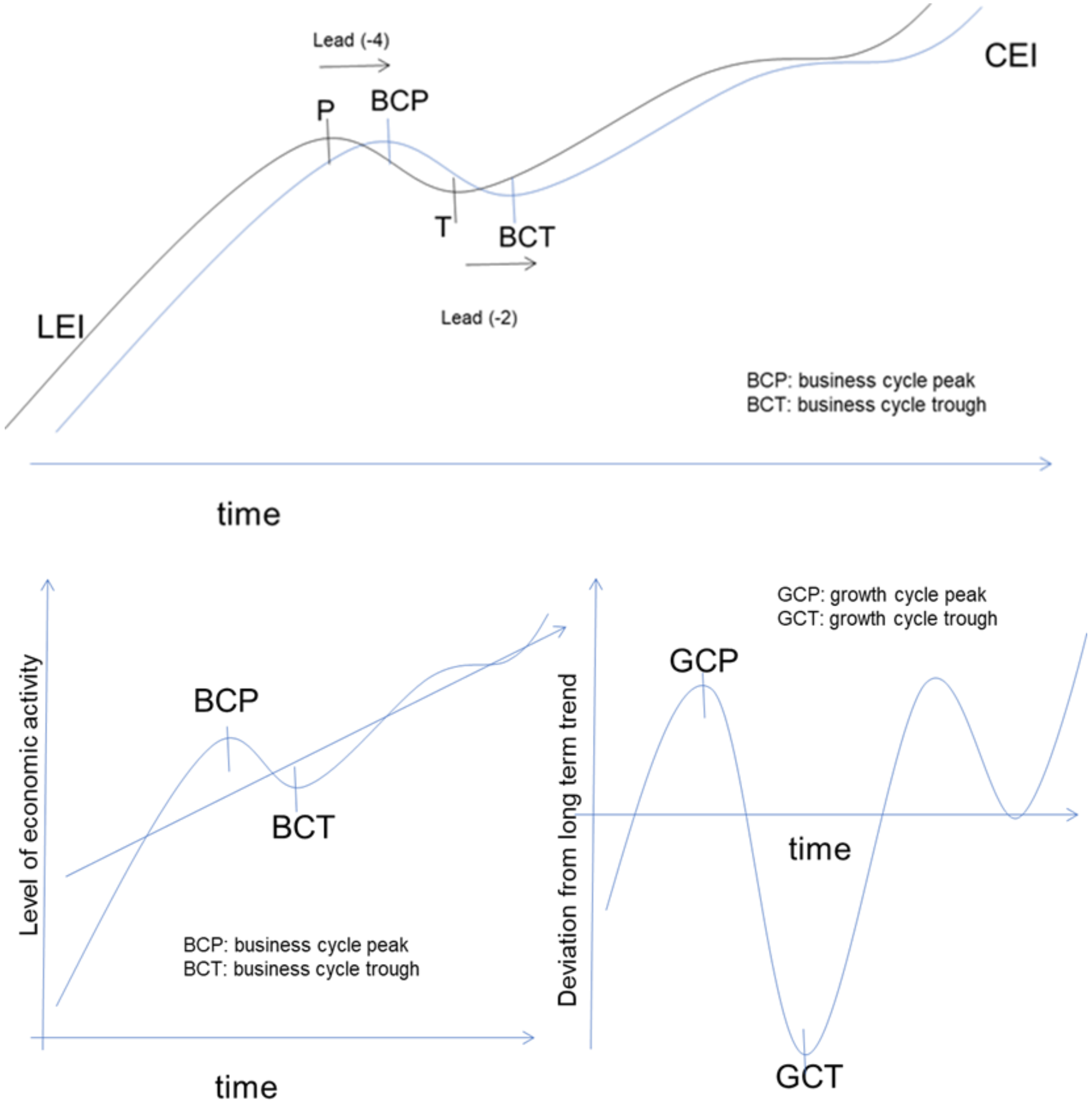
Changes in business and consumer behaviours

Advances in AI and computing infrastructure

Source: https://flyte.org/blog/getting-started-with-large-language-models-key-things-to-know

Turning Points ‘cheatsheet’

- Indicators:
  - Leading, coincident, lagging
  - OECD CLI, Conference Board LEI
- Rules-of-thumb: Sahm, Yield curve
- Econometric models:
  - Nowcasting: AR, VAR, DFM,
  - State changes: Markov Switching, TAR, Logit, Probit
- Machine Learning:
  - Penalised regressions + Classification
  - Neural Networks



Source: The Conference Board, 2022

What makes a good TP detection method?

	Criteria	Description
1	Performance	Accuracy, robustness to revisions
2	Update	Low-cost and rapid
3	Transparency and replicability	Users should have all the necessary information and be able to replicate the indicator
4	Explainability	Intuitive and outcomes easy to understand

Composite Indicators are established turning points predictors

- Composite indicators are designed to lead or coincide with benchmark series, by aggregating country-specific indicators covering key sectors.

- Examples include OECD CLI, Conference Board LEI, EC Surveys.

- Quick to update in normal times, but burdensome in large crises (e.g. pandemic).

- Careful interpretation is needed as can be sensitive to data revisions.

Econometric models vary in strengths and limitations

Nowcasting models

- AR models simple but dependent on revision and past history.
- Vector AR models (VAR) capture interdependencies across series but risks overfitting. Challenging to interpret individual coefficients.
- Dynamic Factor Models (DFM) assume unobserved latent variables drive co-movements of observed time series.
  - Easy to implement but hard to interpret.

State change models

- Logit/Probit models consider binary states, using linear combination of predictors.
  - Sensitive to arbitrary thresholds
- Markov Switching models are flexible to build and combine with other models.
  - Trade-off of accuracy/complexity with interpretability.
- Threshold Auto-Regressive models capture non-linear behaviours well but challenging to apply in real-time and interpret.

Econometric methods

	Indicators	Models					
	Composite Indicators	AR	VAR	DFM	Logit/Probit	MS	TAR
Performance							
Turning point detection		Specific case					
Robustness to revision (end sample)	When 2-sided filters used	No	Depend on inputs	Depend on inputs	Depend on inputs	Depend on inputs	Depend on inputs
Update						One step computational heavy	Computational heavy
Transparence and replicability						More complex	More complex
Explainability			Interpretation of coefficients can be difficult	Factors can be complicated	Interpretation is highly correlated to assumptions	Factors behind regime shifts can be complicated	Can be complex

Machine Learning overlaps traditional statistical models

- ML vs statistical models:
  - Similar roots in statistical theory.
  - Focus on optimising predictive accuracy vs relationship inference.
- Strengths of ML framework:
  - Regularisation mechanisms to prevent overfitting.
  - Cross-validation methods helps optimise model performance.
- Additional benefits (Hirschbuhl 2021):
  - Feature selection and ensemble methods.
  - General forms of non-linearities e.g. extreme shocks.
  - Capturing qualitative data using embedding methods.

ML can outperform standard methods well in data-rich environments

- ML algorithms outperformed MS DFM in detecting TPs (Piger 2019):
  - Faster detection with higher accuracy.
  - Better at trough detection than peak.
- Support Vector Machines and tree/clustering methods performed well in specific contexts; no clear winner across economies and phases.
- Deep Learning (Neural Network) methods promising; potential to apply transfer learning to other countries, but harder to explain.
- Large Language Models could be adapted given sufficient input data and prompt engineering. Explainability a challenge.

ML methods

	Penalised Regressions	Classification methods	Neural networks	LLMs
Performance				
Turning point detection	Context-specific	Context specific, better at troughs than peaks	In rich data environment	Open Q
Robustness to revision (end sample)	Depends on features	Depends on features	Depend on features	Open Q
Update			Choice of architecture	Dependent on retraining
Transparency and replicability			If hyperparameters and architecture are published	Potential commercial paywall
Explainability	Requires nuance to explain penalisation	Some more than others	Can be hard to explain	Hard to explain

Alternative data has a promising future to supplement official sources

- Business and consumer surveys are most popular sources of leading indicators across international statistical/economics organisations
- Administrative data most useful but country-specific and challenging to obtain.
- Alternative quantitative data (public/commercial) promising as early indicators but may lose value and harder to develop narrative.
  - High volume/frequency but may vary widely in coverage, completeness and methodology.
- Improvement in IT infrastructure and ML methods allows text-based data to be handled.
  - Job posting or transitions data can lead labour market indicators.
  - News data can capture uncertainty, sentiment and knowledge graphs to forecast industrial production
  - ML models may differ in performance across datasets in other languages.
- Open source datasets show most promise, considering all criteria.

Alternative Data

Group	Numerical	Text	Image
Performance	Early warning	Can offer early and/or additional insights	Links demonstrated for some industries. Night light has no coverage for certain regions for some months
Country coverage	Source-dependent	Dependent on good quality data, and mature text processing methods for the specific language	Works across country boundaries
Update			
Transparency and replicability	Caveats in collection methodology and completeness	Keyword approaches more transparent than transformer-based methods	Methodology stabilising
Explainability		Keyword approaches more explainable than transformer-based methods	Explainable

(Preliminary) conclusions

- No single method (econometric or ML) outperforms others in turning point detection
- ML could be utilised to extract more timely indicators
- Alternative data offers opportunity to increase set of leading indicators
  - Online job postings for labour markets
  - News data for uncertainty

QUESTIONS