

# TD4 : Construire un instrument d'exposition commerciale (Bartik)

Géoéconomie et Entreprises — M1-M2

Durée estimée : 2h30 | Individuel ou binôme | Remise : avant la séance 10

## Contexte

L'un des défis majeurs de l'économie empirique est d'**identifier un effet causal** quand on ne peut pas réaliser d'expérience contrôlée. Comment mesurer l'impact *causal* de l'invasion de l'Ukraine sur les entreprises européennes, alors que toute l'Europe est touchée ?

La solution utilisée dans la recherche récente est l'**instrument de Bartik** (ou *shift-share instrument*). L'idée est d'exploiter le fait que les pays et secteurs européens n'étaient *pas également exposés* aux importations en provenance de Russie et d'Ukraine *avant* l'invasion. Cette variation pré-existante constitue une source d'identification.

## L'instrument de Bartik : définition

L'exposition Bartik d'un pays  $c$  dans un secteur  $s$  est définie comme :

$$B_{cs} = \sum_{p \in \{RU, UA\}} \omega_{csp}^{\text{pre}} \times \Delta \text{GPR}_p$$

où :

- $\omega_{csp}^{\text{pre}}$  est la **part des importations** du pays  $c$ , secteur  $s$ , en provenance du pays  $p$ , calculée sur la période *pré-choc* (2017–2019)
- $\Delta \text{GPR}_p$  est la **variation du GPR** associée au pays  $p$  (le « choc »)

En pratique, puisque  $\Delta \text{GPR}$  est le même pour tous les pays européens, la variation vient entièrement des **parts d'importation pré-existantes**  $\omega_{csp}^{\text{pre}}$ .

**Intuition** : une entreprise allemande dans un secteur qui importait 15% de ses intrants de Russie est *plus exposée* au choc qu'une entreprise espagnole dans un secteur qui n'importait que 2% de Russie.

## Pourquoi « Bartik » ?

L'instrument porte le nom de Timothy Bartik (1991), qui a proposé cette méthode pour étudier les chocs d'emploi locaux. Il a été récemment formalisé par Goldsmith-Pinkham, Sorkin & Swift (2020, *AER*). La clé de la validité est que les parts pré-existantes ( $\omega^{\text{pre}}$ ) ne sont *pas affectées par le choc* — c'est la structure historique du commerce qui fournit l'identification.

## Prérequis

- Python 3.8+ avec pandas, matplotlib, seaborn, numpy
- Installation : `pip install pandas matplotlib seaborn numpy`

## Données

Fichier	Description	Obs.
trade_bilateral.csv	Importations bilatérales, 8 pays × 6 secteurs	~1 000
gpr_annual.csv	GPR annuel global (2015–2025)	11

### Variables dans trade\_bilateral.csv :

Variable	Description
importer	Code pays importateur (DE, FR, IT, ES, NL, BE, AT, PT)
exporter	Code pays exportateur (RU, UA, CN, US, et autres)
sector	Secteur NACE simplifié (6 secteurs)
year	Année (2015–2023)
import_value	Valeur des importations (millions USD)

### Secteurs :

Code	Secteur
ENERGY	Énergie (gaz, pétrole, charbon)
METALS	Métaux et minéraux
AGRI	Agroalimentaire (céréales, oléagineux)
CHEM	Chimie et pharmacie
MACH	Machines et équipements
OTHER	Autres manufactures

## Exercice 1 : Explorer les données commerciales (30 min)

1. Charger le fichier `trade_bilateral.csv`
2. Pour chaque pays importateur, calculer la **part des importations en provenance de Russie** (tous secteurs confondus) sur 2017–2019
3. Créer un graphique en barres montrant cette part par pays
4. Identifier le pays le plus et le moins exposé à la Russie
5. Répéter l'analyse par **secteur** (tous pays confondus)

*Question : Quels sont les 2 secteurs les plus dépendants de la Russie ? Est-ce cohérent avec ce que vous savez de la structure commerciale européenne ?*

## Exercice 2 : Construire l'instrument Bartik (45 min)

1. Calculer les **parts d'importation pré-choc**  $\omega_{cs,RU}^{\text{pre}}$  pour chaque paire (pays, secteur) sur 2017–2019 :

$$\omega_{cs,RU}^{\text{pre}} = \frac{\text{Imports}_{cs,RU}^{2017-2019}}{\text{Imports}_{cs,\text{total}}^{2017-2019}}$$

2. Créer une **matrice pays × secteur** de ces parts (tableau ou heatmap)
3. Calculer l'exposition Bartik simplifiée pour chaque pays :

$$B_c = \frac{1}{S} \sum_s \omega_{cs,RU}^{\text{pre}}$$

où  $S$  est le nombre de secteurs

4. Créer un graphique en barres de  $B_c$  par pays
5. Vérifier que la **variation** de  $B_c$  entre pays est substantielle (ratio max/min)

*Question : Pourquoi est-il important que la variation de l'instrument soit suffisamment grande ? Que se passerait-il si tous les pays avaient la même exposition ?*

### Exercice 3 : Avant/après Ukraine (30 min)

1. Calculer les parts d'importation en provenance de Russie **avant** (2017–2019) et **après** (2022–2023) l'invasion
2. Créer un graphique montrant la **chute** des importations russes par pays
3. Calculer la **corrélation** entre l'exposition pré-choc ( $B_c$ ) et la baisse des importations russes
4. Tester si les pays plus exposés ont davantage réduit leurs importations russes

*Question : La corrélation est-elle positive, négative ou nulle ? Est-ce que cela conforte ou affaiblit la validité de l'instrument ?*

### Exercice 4 : Discussion de validité (30 min)

Cet exercice est conceptuel. Répondez aux questions suivantes en vous appuyant sur le cours (Séance 9) et sur vos résultats :

1. **Hypothèse d'exogénéité** : Les parts d'importation pré-choc  $\omega^{\text{pre}}$  sont-elles vraiment exogènes au choc ? Quels arguments pour et contre ?
2. **Pre-trends** : Si les importations de Russie étaient déjà en baisse avant 2022 (à cause de Crimee 2014 ou des sanctions 2014–2021), cela pose-t-il un problème ?
3. **Exclusion restriction** : L'instrument  $B_c$  n'affecte-t-il les entreprises *que* via le canal commercial ? Ou y a-t-il d'autres canaux (énergie, finance, incertitude) qui corrélerent avec l'exposition ?
4. **GPSS diagnostics** : Goldsmith-Pinkham et al. (2020) recommandent de vérifier les « poids de Rotemberg » (la contribution de chaque part à l'estimation). Pourquoi est-il problématique qu'un seul pays domine ?

*Ces questions n'ont pas de réponse unique. L'objectif est de montrer votre compréhension des forces et faiblesses de la stratégie d'identification.*

### Bonus : pour aller plus loin

1. **Heatmap complète** : créer une heatmap pays  $\times$  secteur de la variation des importations russes (pré vs post)
2. **Indice HHI** : calculer l'indice de Herfindahl-Hirschman de concentration des importations par pays et secteur. Les secteurs les plus concentrés sont-ils les plus vulnérables ?
3. **Diversification** : pour chaque paire (pays, secteur), calculer le changement de la part russe entre 2019 et 2023. Qui a le plus diversifié ses sources ?

### Rendu

- Un fichier Jupyter Notebook (.ipynb) ou un script Python (.py) commenté
- 5+ figures (parts par pays, heatmap, barres Bartik, chute importations, corrélation)
- Réponses aux questions (exercices 1–4), en particulier la discussion de validité (exercice 4)
- Dépôt sur Moodle avant le début de la séance 10

## Critères d'évaluation

Critère	Points
Code fonctionnel et lisible	25%
Qualité des figures (labels, titre, source, légende)	20%
Construction correcte de l'instrument Bartik	25%
Discussion de validité (exercice 4)	20%
Présentation générale	10%

## Ressources

- Goldsmith-Pinkham, P., Sorkin, I. & Swift, H. (2020). « Bartik Instruments : What, When, Why, and How ». *American Economic Review*, 110(8), 2586–2624.
- Bartik, T. (1991). *Who Benefits from State and Local Economic Development Policies?* Upjohn Institute.
- Séance 9 : Stratégie empirique et résultats — identification par Bartik.
- Base de données BACI (CEPII) : [http://www.cepii.fr/CEPII/en/bdd\\_modele/bdd\\_modele\\_item.asp?id=37](http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37)

*Un bon instrument d'identification est comme une clé : il n'ouvre qu'une seule porte. Si votre instrument ouvre plusieurs portes, c'est un passe-partout — et il ne vous apprend rien.*