

Politiques du marché du travail fondées sur des données probantes - Fondamentaux méthodologiques et innovants Exemple

Prof. Dr. Marco Caliendo
(University of Potsdam and Evaluation Office)
`caliendo@uni-potsdam.de`
`www.caliendo.de`
`www.evaluation-office.de`

WAPES World Congress
April 28, 2022

Évaluation de l'impact contrefactuel

- **Objectif** : Estimer l'impact causal d'une certaine politique sur les « unités » touchées!
- Question centrale à laquelle répondre :
« **Que se serait-il passé si les unités touchées n'avaient pas reçu le traitement?** » (résultat contrefactuel)
- **Effet causal** : Comparaison des résultats observés avec situation **contrefactuelle**.
- **Problème fondamental de l'évaluation** : ce contrefactuel n'est jamais observé (pour la même unité en même temps)!
⇒ Par conséquent, nous devons trouver **un bon échantillon** (proxy) dans un groupe de comparaison!

Biais de sélection

- Le principal problème auquel nous sommes confrontés avec les « approches standard » (p. ex., les comparaisons avant-après/transversales) est que l'assignation au traitement et au groupe témoin **n'est pas aléatoire**.
- Les participants et les non-participants peuvent **différer même en l'absence du programme**.



Treatment Group



Comparison Group

- Par conséquent, une simple comparaison (moyenne) n'est pas significative en raison du **biais de sélection**.

Résoudre le problème de sélection

– Il existe une variété de méthodes bien établies pour surmonter le biais de sélection. **Notre focus aujourd'hui :**

Méthodes expérimentales



Ill. Niklas Elmehed. © Nobel Media.
Abhijit Banerjee
Prize share: 1/3



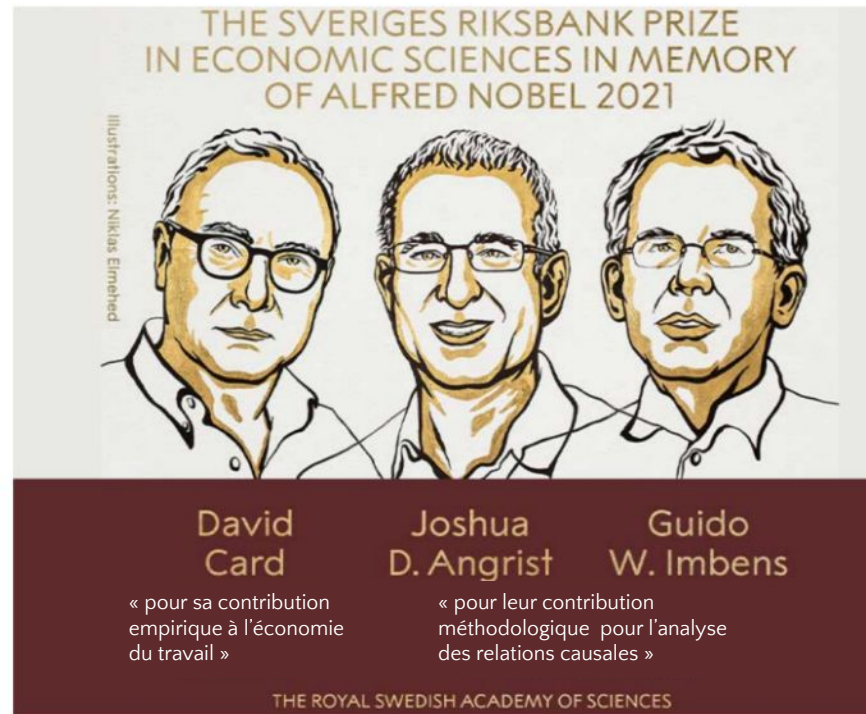
Ill. Niklas Elmehed. © Nobel Media.
Esther Duflo
Prize share: 1/3



Ill. Niklas Elmehed. © Nobel Media.
Michael Kremer
Prize share: 1/3

Le prix Sveriges Riskbank en Sciences Economiques à la mémoire d'Alfred Nobel 2019 a été décerné conjointement à Abhijit Banerjee, Esther Duflo et Michael Kremer « pour leur approche expérimentale de la réduction de la pauvreté dans le monde ».

Méthodes quasi/non expérimentales



Sommaire

1 Introduction

2 Cadre d'évaluation

3 Méthodes d'évaluation

1 Expériences

2 Matching

3 Différence-dans-les-différences (DiD)

4 Design de Discontinuité de Régression (DDR)

4 Conclusion

Évaluation du programme - Un scénario dans un monde idéal(2)

- Dans un monde idéal, l'évaluateur est déjà impliqué aux premières étapes de la conception du programme et a une influence sur les données recueillies pour une évaluation ultérieure.
- Ces étapes sont :
 1. Définir les objectifs du programme
 - Développer une théorie du changement
 - Concevoir le programme
 - Mettre en œuvre et collecter des données de référence
 - Recueillir des données sur les résultats finaux
 - Évaluation de l'impact contrefactuel
- L'évaluation des processus (accent mis sur la mise en œuvre et l'exploitation du programme) et l'évaluation de l'impact devraient être considérées comme complémentaires.
- Nous pouvons utiliser les informations recueillies lors de l'évaluation des processus pour choisir parmi d'autres critères d'évaluation.

Évaluation du programme - Un scénario dans un monde idéal(2)

Questions importantes auxquelles il convient déjà de répondre au stade de la conception:

- **Objectifs et mesure du succès :**

- Quels sont les effets escomptés du programme?
- Comment mesure-t-on le succès du programme?

- **Théorie du changement:**

- Quelle est la séquence des événements qui mène aux résultats observés?
- Quels sont les différents paramètres qui contribuent au succès du programme?

- **Stratégie empirique :**

- Quel type de méthodologie d'évaluation doit être poursuivi?
- Comment les données nécessaires seront-elles recueillies?
- Comment distinguer quels mécanismes théoriques sont les plus importants ?

⇒ Dans un monde idéal, les évaluateurs disposent de suffisamment de **temps**, de **budget** et de **données de haute qualité**.

Évaluation du programme - Le scénario du monde réel

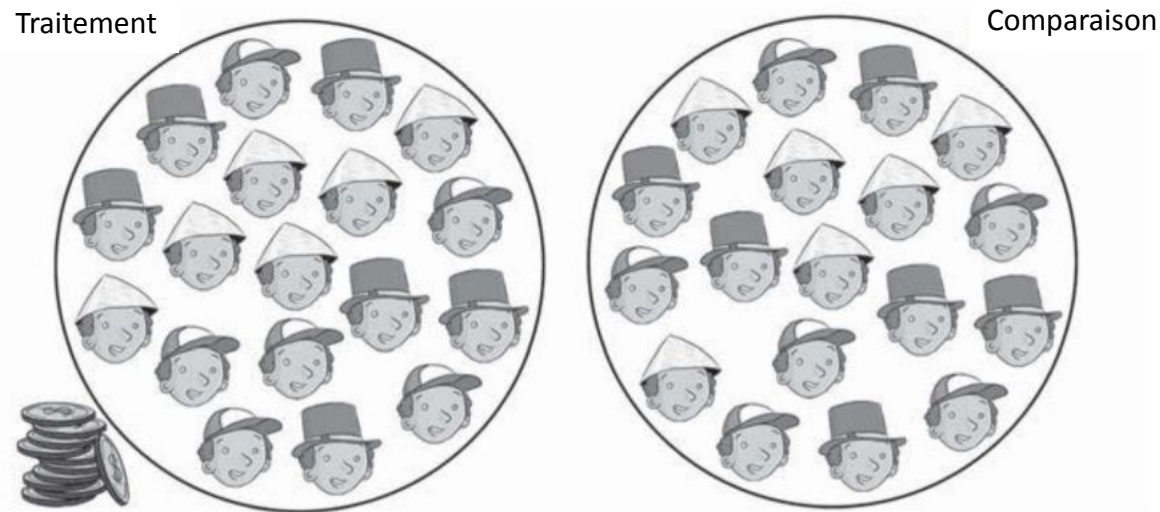
– Cependant, dans le **monde réel**, les évaluations sont souvent effectuées dans des circonstances moins qu’optimales (« **évaluations à la rigueur** »):

Les contraintes dans lesquelles les évaluations doivent être effectuées			
Temps	Budget	Data	Scénario typique
X			L'évaluateur est sollicité tardivement avec un délai restreint
	X		Difficultés à recueillir les données de l'enquête
		X	Pas de données de base disponibles, sujet sensible avec une collecte de données difficile
X	X		Des données secondaires sont disponibles mais peu de temps pour les analyser
		X	Peu de temps et aucune donnée n'ont été collectées conception de l'enquête limitée en raison de contraintes de temps
	X	X	L'évaluateur est appelé en retard, la date limite n'est pas un problème Pas d'accès aux données de base, le budget est serré
X	X	X	L'évaluateur est appelé en retard avec un délai serré et un budget serré, aucune donnée de base et aucun groupe témoin n'a été identifié

Source: Bamberger et. al (2004)

Expériences

- **Les expériences** attribuent des unités de la population éligible de **manière aléatoire**.
- Cela garantit que la participation n'est pas liée aux caractéristiques des unités et équilibre les échantillons dans les caractéristiques observées et non observées.



Source: Evaluation in Practice

- Par conséquent, les différences de résultats observées entre les deux groupes peuvent être attribuées **uniquement** au traitement!
- **Estimateur** : Différences moyennes transversales simples dans le résultat Y.

1. Introduction	2. Cadre d'évaluation	3.1 Expériences	3.2 Matching	3.3 DiD	3.4 RDD	4. Conclusions	Appendix
●●●●	●●●	●●	●●●	●●●●●	●●●	●●	●●●●●●●

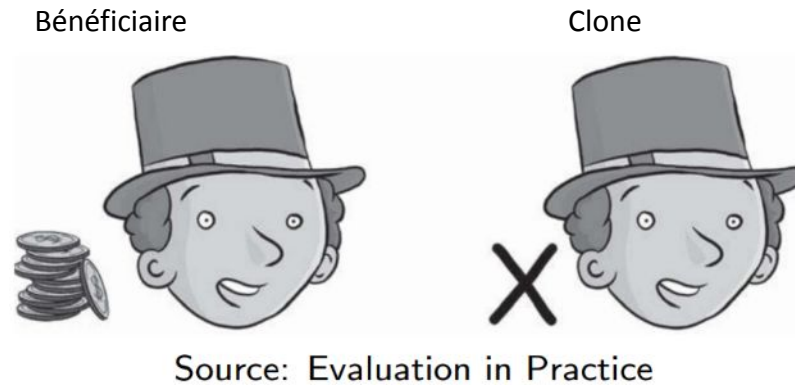
Exemple d'expérience : Déplacement d'emplois

Crépon et al. (2013): Les politiques du marché du travail ont-elles des effets de déplacement ? Preuves d'une expérience randomisée en grappes

- **Question de recherche** : Quel est l'effet direct et indirect (déplacement) de l'aide au placement sur les résultats sur le marché du travail des jeunes demandeurs d'emploi instruits en France ?
- **Traitement** : Participation à un programme de placement.
- **Données** : Données administratives et d'enquête, recueillies en 2007-2010. N ≈ 22 000.
- **Méthode** : Expérience de terrain contrôlée randomisée. Les jeunes sans emploi et instruits ont été assignés au hasard à un groupe de traitement ou de contrôle.
- **Résultats** : Il y a un effet positif pour les jeunes traités sur la recherche d'emploi, tandis que les jeunes non traités sont moins susceptibles de trouver un emploi (effet de déplacement).

Appariement (Matching)

- **Idée** : Choisissez pour chaque participant, un (ou plusieurs) **jumeaux statistiques** parmi l'échantillon de non-participants.
- Ils doivent être identiques pour **toutes les caractéristiques pertinentes**! Il s'agit d'une exigence très forte qui nécessite des données informatives.



- Semblable à une expérience, cela conduit à un échantillon équilibré.
- **Estimateur** : Différences des moyennes transversales simples dans le résultat sur **l'échantillon apparié**.
- **Malédiction de la dimensionnalité**: Si le nombre de caractéristiques pertinentes est important, il peut être très difficile de trouver une **correspondance exacte**!
- **Une solution** : l'appariement « propension-score » résume toutes les informations dans un seul indice et choisir le non-participant le plus proche au regard de cet indice.

De meilleures données aident beaucoup!

- Mettre en œuvre une approche d'appariement crédible n'est pas facile.

De meilleures données aident beaucoup!

- Souvent, les estimations peuvent être améliorées en combinant plusieurs sources de données :
 - Les données au niveau des individus et des entreprises sont souvent disponibles, à faible coût, dans des enregistrements administratifs (p. ex., par l'entremise des agences pour l'emploi).
 - Les données régionales/nationales sont fournies par des données (inter)nationales **des agences de statistique.**
- **Exemples :**
 - L'ensemble de données d'évaluation IZA combine des données administratives et d'enquête permettant d'enrichir les données de l'administration avec des informations sur «caractéristiques habituellement non observées »
 - DellaVigna et al. (2022) mènent une enquête par SMS à deux reprises une semaine pendant quatre mois.

Exemple d'appariement : Subventions au démarrage

Caliendo/Künn/Weißenberger (2016): Traits de personnalité et évaluation de la subvention de démarrage

- **Question de recherche** : Les subventions au démarrage pour les chômeurs constituent-elles un programme actif efficace sur le marché du travail? Et les traits de personnalité omis constituent-ils une menace pour la fiabilité des estimations correspondantes?
- **Traitement**: Les chômeurs désireux de créer une entreprise obtiennent des transferts mensuels pour une durée maximale de 15 mois.
- **Données** : Combinaison de données administratives et d'enquêtes. $N \approx 1\ 300$.
- **Source du biais de sélection** : Les participants s'auto-sélectionnent dans le programme; les participants diffèrent par leurs caractéristiques des non-participants!
- **Méthode** : Appariement des participants et des non-participants en fonction d'un large éventail de caractéristiques et de résultats avant le traitement.
- **Résultats** :
 - Effets positifs sur les probabilités d'emploi et les revenus.
 - Les résultats sont renforcés par la prise en compte de traits de personnalité généralement non observés!

Différence dans les Différences (1)

- Les configurations de **différence dans les différences** (DiD) exploitent souvent une sorte d'« expérience naturelle » qui se produit en raison d'un changement de politique, lorsqu'un groupe d'unités est touché par le traitement alors que l'autre groupe n'est pas affecté.
 - **Par exemple** : un État augmente le salaire minimum, mais l'État voisin ne le fait pas.
- **Important**: DiD suppose des tendances temporelles parallèles (TTP) pour le traitement et le contrôle du groupe témoin sans traitement, et, permet différents niveaux de prétraitement (« biais de base »).
- **Validité des TTP** :
 - L'inspection de la similitude des tendances de prétraitement fournit quelques indication de la probabilité que l'hypothèse des TTP soit valable.
 - Des tendances de prétraitement significativement différentes jettent de sérieux doutes sur la fiabilité des estimations

Différence dans les Différences (2)

- **Intuition de l'Estimateur DiD** : Combiner des estimations avant-après pour le traitement et le groupe témoin.
 - En comparant les changements au sein des groupes, nous contrôlons implicitement les **facteurs non observés constants dans le temps**.
 - En comparant ces changements entre les groupes, nous contrôlons également les **tendances temporelles des résultats**.
- **Estimateur** : Comparez les changements dans le groupe de traitement au fil du temps avec les changements dans le groupe de comparaison au fil du temps.

Exemple DiD - Salaire minimum (1)

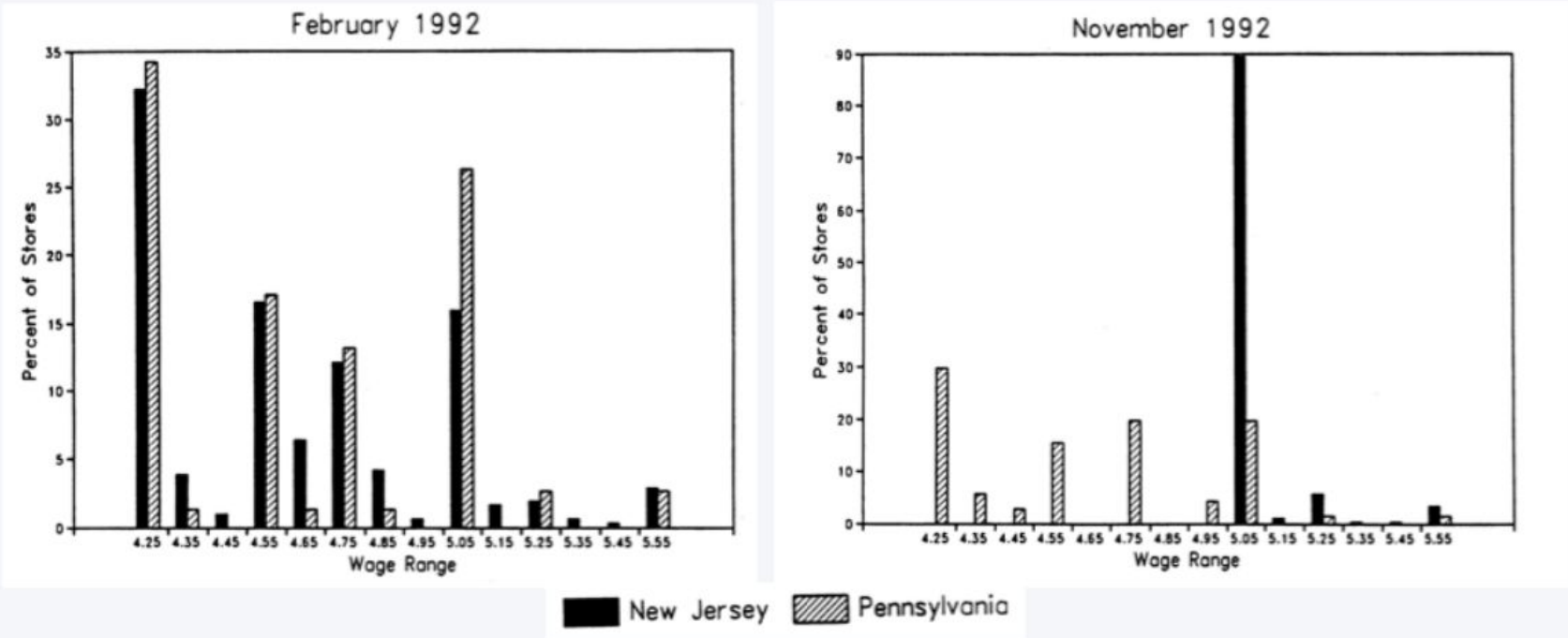
Card/Krüger (1994): Salaire minimum et emploi: étude de cas de l'industrie de la restauration rapide dans le New Jersey et la Pennsylvanie

- **Question de recherche** : Incidence de l'augmentation du salaire minimum sur l'emploi à faible salaire?
- **Traitement**: Augmentation du salaire minimum de 4,25 \$ à 5,05 \$ l'heure dans le New Jersey en avril 1992.
- **Données** : Données d'enquête sur les salaires et l'emploi pour N = 410 restaurants de restauration rapide dans le New Jersey et la Pennsylvanie.
- **Source de biais de sélection**: Les restaurants non affectés dans le New Jersey peuvent servir à différents clients et offrir des repas plus chers.
- **Méthode**: Comparez l'évolution de l'emploi à temps plein dans les restaurants de restauration rapide dans le New Jersey et l'État voisin PA.

Exemple DiD - Salaire minimum (2)

Card/Krüger (1994): Salaire minimum et emploi: étude de cas de l'industrie de la restauration rapide dans le New Jersey et la Pennsylvanie

— Comparaison descriptive des salaires avant et après traitement.



Exemple DiD - Salaire minimum (2)

Card/Krüger (1994): Salaire minimum et emploi: étude de cas de l'industrie de la restauration rapide dans le New Jersey et la Pennsylvanie

– Le calcul des rendements moyens de l'échantillon (s.e. entre parenthèses) :

Variable	Struct. par Etat		
	PA (i)	NJ (ii)	Différence, NJ – PA (iii)
1. ETP employés avant, toutes observations disponibles	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)
2. ETP employés après, toutes observations disponibles	21.17 (0.94)	21.03 (0.52)	-0.14 (1.07)
3. Variation de l'emploi moyen dans les ETP	-2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

Source: Card/Krueger (1994), p. 780

Design de Discontinuité de Régression (DDR)

- De nombreux programmes fonctionnent avec un certain **seuil d'admissibilité** en ce qui concerne certains indices.
- Par exemple, les travailleurs au-delà d'un certain âge reçoivent des **allocations de chômage** pour une durée plus longue.
- Le DDR (brut) compare les résultats moyens des unités juste en dessous et juste au-dessus du seuil.
- La différence donne une estimation des **effets moyens locaux** du traitement du programme pour les personnes **au seuil**.

Exemple DDR: Prestations de chômage (1)

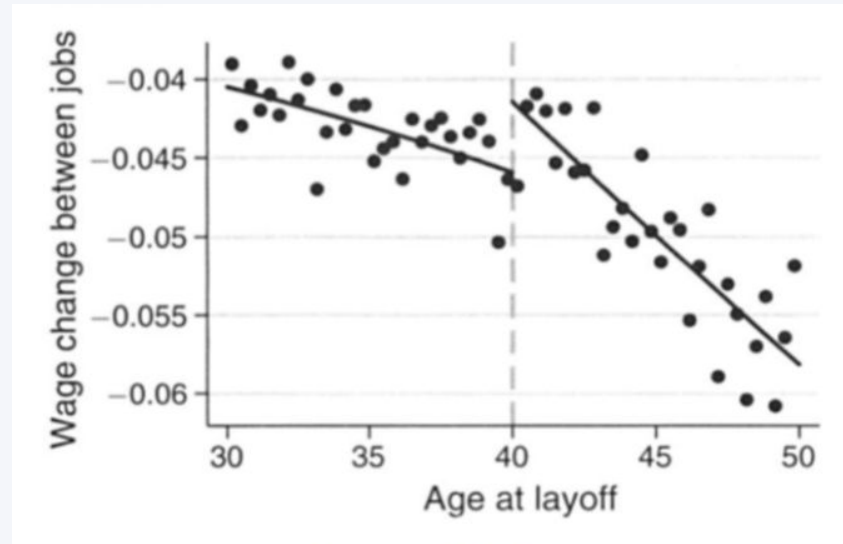
Nekoei/Weber (2017): L'extension des allocations de chômage améliore-t-elle la qualité de l'emploi?

- **Question de recherche** : Impact de l'allongement de la durée des allocations de chômage sur les salaires post-chômage ?
- **Traitement**: Le 1er août 1989, le gouvernement autrichien a augmenté la durée potentielle des allocations de chômage de 30 à 39 semaines pour les personnes de 40 ans et plus au moment de la cessation d'emploi.
- **Données**: Données administratives sur les salaires et l'emploi pour un échantillon de $N \approx 1,7$ million de départs involontaires.
- **Source du biais de sélection**: Les chômeurs non affectés peuvent avoir moins d'expérience du marché du travail et des niveaux d'éducation différents.
- **Méthode** : Comparer les salaires de réemploi des personnes légèrement plus jeunes et âgées de plus de 40 ans lorsqu'elles perdent leur emploi.

Exemple DDR: Prestations de chômage (2)

- Graphiquement, l'effet du traitement est égal au saut vertical dans la variable de résultat au seuil.

Nekoei/Weber (2017): L'extension des allocations de chômage améliore-t-elle la qualité de l'emploi?



Source: Fig. 3

- **Interprétation:** L'extension de l'éligibilité à l'AC augmente les salaires de réemploi d'environ 0,5% pour les personnes au seuil.

Conclusions

- Pour assurer une **évaluation réussie du programme** ...
 - ... **Impliquer l'évaluateur** le plus tôt possible.
... **planifier la méthode d'évaluation** lors de la conception du projet.
- **Important** : De meilleures données aident beaucoup !
 - La **combinaison de différentes sources de données** (p. ex. administratives et données d'enquête) peuvent être utiles dans de nombreuses situations (mais peuvent également prendre du temps)! Il peut être utile de penser à la **collecte de nouvelles données** spécifiquement pour l'évaluation du programme.

« Vous pouvez tout faire. Mais vous ne pouvez pas tout faire et vous ne peut certainement pas tout faire en même temps.

Merci pour votre attention!