

Mesurer la réponse de la consommation à des chocs de revenus : l'apport des données bancaires à haute fréquence

Odran Bonnet¹ François Le Grand² Tom Olivia³ Xavier Ragot⁴ Lionel
Wilner¹

¹Insee-Crest

²EM Lyon, ETH Zürich

³Insee

⁴Sciences Po, CNRS, OFCE

Journées de Méthodologie Statistique - 30 Mars 2022

Plan

Motivation

Forces et faiblesses des données bancaires

Résultats empiriques

Chocs transitoires

Chocs permanents

Conclusion

Questions normatives de politique économique

Quelle réaction de la consommation des ménages à des chocs de revenus ?

Objectif : éclairer l'élaboration des politiques publiques de type "transferts directs aux ménages"

1. qui doit bénéficier des aides ? meilleur ciblage : caractérisation fine des populations qui réagissent davantage aux politiques de relance
2. combien ces personnes doivent-elles recevoir ? détermination des montants optimaux : estimation des effets revenus

Contribution : estimation de la MPC à partir de différents types de chocs de revenus

Nous examinons deux catégories de chocs à travers différentes expériences :

- Chocs **transitoires** : transferts directs (primes Covid) et paiements réguliers (jour de paye, allocation de rentrée scolaire)
- Chocs **permanents** : pertes d'emploi (choc **néгатif**) et hausses de salaires (choc **positif**)

Revue de littérature

En économie publique, une littérature est consacrée à l'étude de la sensibilité excessive de la consommation (ESC) au revenu : Shea (1995), Angeletos et al. (2001), Gelman et al. (2014), etc.

Différentes estimations de propensions marginales à consommer (MPC) sont disponibles, selon les pays, ménages, périodes considérés :

- MPC trimestrielles : comprises entre 0,5 et 0,75 (Broda et Parker 2014) - USA, stimulus fiscal (crise de 2008)
- MPC mensuelles sur l'aide alimentaire : de l'ordre de 0,5-0,6 (Hastings et Shapiro 2018) - USA, 2004-2016
- MPC trimestrielles sur remise d'impôt : effet hétérogène et décroissant selon les liquidités disponibles, Gelman (2021) - USA, 2012-2016. Baugh et al. (2018), asymétrie de la MPC selon le signe du choc (paiement contre remboursement d'impôt)
- MPC à 5 semaines sur l'obtention de l'Allocation de Rentrée Scolaire (ARS) : 0,38 (Landais et al. 2021) - France, 2020
- Élasticité de la consommation au revenu suite à des chocs de salaires : entre 0,3 et 0,38 (Baker 2018) - USA, 2007-2013
- MPC à 24 mois suite à une perte d'emploi : environ 0.3 - Danemark, 2009-2016 (Andersen et al., 2021)

L'effet de la perception de la paie (un revenu régulier et prévisible), mesuré au niveau journalier, est l'objet d'une littérature spécifique (Stephens Jr. 2003, 2006, Olafsson et Pagel 2018).

Plan

Motivation

Forces et faiblesses des données bancaires

Résultats empiriques

Chocs transitoires

Chocs permanents

Conclusion

Une nouvelle source de données à très haute fréquence

Les données bancaires : un éclairage inédit dans des délais très resserrés

- opérations par cartes, virements, prélèvements, chèques, retraits, soldes mensuels, etc.
- suivi individuel, longitudinal et à haute fréquence (journalière)
- deux panels anonymisés d'environ 300 000 clients (Crédit Mutuel Alliance Fédérale et La Banque Postale)
- données mises à jour très régulièrement, un suivi quasiment en temps réel de 3 variables clés (consommation, revenus, épargne)
- possibilité de suivre une sous-population précise (ex : bénéficiaires du RSA, nouveaux chômeurs, . . .)
- deux exploitations conjoncturelles dans le cadre de la crise sanitaire :
 1. Note de conjoncture (mars 2021) : un retour sur un an de crise sanitaire
graphique patrimoine
 2. Insee analyses (novembre 2021) : un focus sur les populations précaires (bas revenus, RSA) graphique RSA

Caractéristiques des données

Deux bases de données, deux réseaux bancaires :

- *Crédit Mutuel Alliance Fédérale (CM)* : 300 000 clients échantillonnés parmi 4,2 millions de clients, suivis en panel entre février 2019 et février 2022
- *La Banque Postale (LBP)* : 300 000 clients échantillonnés parmi 11 millions de clients, suivis en panel entre janvier 2019 et novembre 2021
- **Points communs** : données au niveau du ménage, présence de quelques variables socio-démographiques (âge, sexe, département, proxy de la PCS, statut marital, . . .), suivi des soldes au niveau mensuel et des opérations au niveau journalier
- **Différences** :
 - Champ : France métropolitaine hors Corse (CM) vs France entière (LBP)
 - Labellisation des opérations : paiements par carte et prélèvements (CM) vs aucune (LBP)

Concepts clés

Dépenses de consommation :

- Somme des dépenses par carte et des retraits en espèces.
- Hors virements sortants et chèques.

Revenus :

- Repérage délicat sans labellisation des virements entrants.
- « Revenu » : somme des virements entrants inférieurs à 40 000 € et sans centimes.

Patrimoine financier : Somme du patrimoine liquide brut et du patrimoine illiquide brut :

- Patrimoine liquide brut : somme des soldes des comptes courants et des différents livrets d'épargne (livret A, livret jeune, CEL, PEL, ...).
- Patrimoine illiquide brut : somme des soldes des assurances vies et des comptes titres.

Crédits :

- A la consommation et immobiliers (contractés dans la banque).
- Empiriquement, utilisation de la présence de crédits seulement.

Filtrage des données

Concept de client actif : Les échantillons de 300 000 clients ont été tirés dans la clientèle de chaque banque (stratification : âge quinquennal × département).

Conservation des comptes bancaires actifs :

- Le ménage doit contenir au moins un adulte
- Il y a toujours au moins 150 € de dépenses et de revenus sur 3 mois glissants sur le compte
- Le compte doit être présent sur toute la période

Les échantillons finaux contiennent environ 170 000 observations pour le CM et 220 000 observations pour LBP.

Représentativité des données

Pour chaque échantillon, nous effectuons un calage sur marges (variables : âge, sexe et département) pour être représentatif du champ concerné en 2019.

Représentativité :

- Bonne reproduction des dépenses de consommation observées dans l'enquête *Budget des Familles* le long de la distribution des revenus.
- Les clients du CM sont en moyenne plus riches en patrimoine et en revenus que la population française.

[graphique conso BDF](#)

[tableau stats descs](#)

Exhaustivité :

- Au niveau macro, des niveaux de patrimoine liquide bruts proches de ceux mesurés par La Banque de France.
- Pour les revenus, plus de volatilité au CM que dans les données de comptabilité nationale.

[graphique patrimoine liquide BDF](#)

[graphique revenus compta nat](#)

Limites des données

- La représentativité :
 - la clientèle n'est pas représentative de la population française, malgré le calage sur marges
 - deux raisons principales :
 1. la clientèle d'un groupe bancaire est marquée socialement (plus aisée au CM, moins aisée à LBP, en moyenne)
 2. nos panels n'ont pas d'entrants, le vieillissement du panel conduit en moyenne à une hausse des revenus et des avoirs sur les comptes.
- La multi-bancarisation :
 - un ménage peut avoir plusieurs comptes dans différentes institutions financières : vision partielle du patrimoine financier
- Virements et chèques sans libellés :
 - difficile d'identifier les revenus.
- Absence des codes MCC à LBP :
 - impossible de décomposer les dépenses de consommation par grande catégorie.

Plan

Motivation

Forces et faiblesses des données bancaires

Résultats empiriques

Chocs transitoires

Chocs permanents

Conclusion

L'aide exceptionnelle covid de novembre 2020 (1/2)

Contexte : Versements d'aides exceptionnelles dites « covid » en mai et novembre 2020. Afin de purger l'effet « confinement strict », on se concentre sur le second versement.

Champ : Bénéficiaires de minimas sociaux (RSA, RSO, ASS) et d'aides au logement (APL). Le montant reçu est de 150 € (+ bonus de 100 € par enfant). Repérage par le montant et la date de versement.

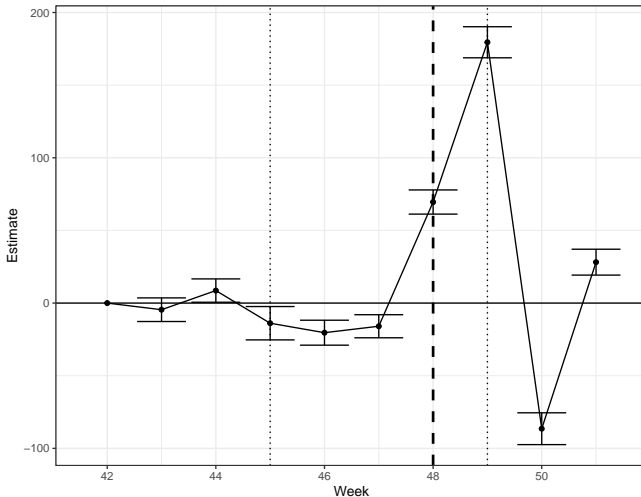
Stratégie d'identification : Appariement par score de propension entre les individus traités en 2020 et des individus similaires en 2019. \simeq 5 400 traités et 5 400 contrôles.

Estimation : Modélisation en différences de différences : « event study » au niveau semaine \times individu + modèle en différences i.e différence de conso totale après-avant régressée sur le montant de l'aide reçue.

$$C_{i,t} = \alpha_0 + \alpha_T T_i + \sum_{k=43}^{51} \beta_k \mathbb{1}_{\{t=k\}} + \sum_{k=43}^{51} \gamma_k \mathbb{1}_{\{t=k\}} T_i + \delta X_i + \varepsilon_{i,t} \quad (1)$$

$$\Delta C_i = \beta_0 + \beta_T \Delta T_i + \varepsilon_i \quad (2)$$

L'aide exceptionnelle covid de novembre 2020 (2/2)



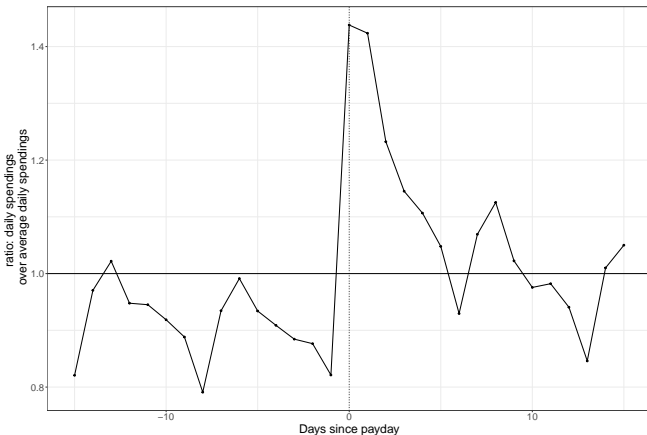
Résultats :

- MPC de 0.89 (écart-type : 0.04).
- Effet décroissant avec le patrimoine liquide.

Effet du jour de paie (1/3)

Contexte :

- Estimation du surcroît de dépenses de consommation (« payday effect ») le jour de paie relativement à un jour « moyen ».
- Exploitation de la granularité fine des données bancaires (variabilités individuelles et temporelles des dates de paie).



Effet du jour de paie (2/3)

Concepts et champ :

1. Mesure de la paie :

- virement entrant le plus élevé du mois (si supérieur à 80% des virements entrants dans une fenêtre de 29 jours)

2. Construction d'un indicateur statistique du « payday effect » :

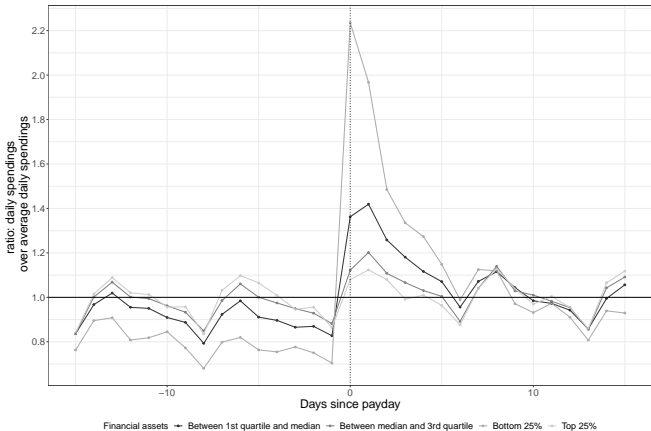
- $x_{i,t} = (29 \times c_{i,t}) / \sum_{k=-14}^{+14} c_{i,k} = c_{i,t} / c_{i,\cdot}$
- par construction, $x_{i,\cdot} = 1$
- un ratio supérieur à 1 signifie que la consommation un jour donné est plus élevée que la consommation moyenne du « mois » considéré

détails

Stratégie d'identification et estimation : modèle en panel au niveau journalier, inspiré de Olafsson et Pagel (2018)

$$x_{i,t} = \sum_{k=-7}^{+7} \beta^k \text{Paid}_{i,t+k} + \delta_{dow} + \phi_{wom} + \psi_m + \zeta_y + \eta_i + \varepsilon_{i,t} \quad (3)$$

Effet du jour de paie (3/3)



Résultats :

- $\beta^0 = 0.43$, dépenses de consommation 1,4 fois plus haute le jour de paie qu'un jour « classique » en moyenne.
- Effet plus fort pour le premier quartile de patrimoine liquide.

plus de résultats

Les pertes d'emploi (1/2)

Contexte : La littérature (Andersen et al., 2021 et Ganong et Noel, 2019) montre qu'une partie du choc de revenus à l'entrée au chômage est absorbée par une baisse de la consommation.

Concepts et champ :

- Pas d'observation directe de la perte emploi.
- En revanche, on observe la perception d'allocations de pôle emploi.
- Plusieurs restrictions conservatrices nous conduisent à garder $\simeq 2\,900$ individus dans notre sous-échantillon de chômeurs.

détails

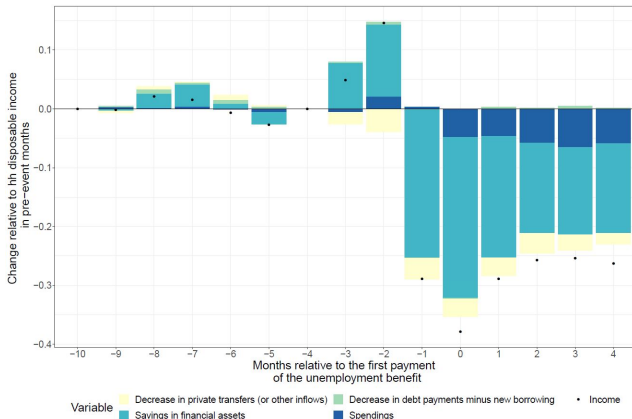
Stratégie d'identification : Différences de timing des périodes de chômage entre les individus (cf. Andersen et al., 2021).

Estimation : Modèle « event study » au niveau mois \times individu. Effets fixes individuels et temporels (« two-way fixed effects »). Intérêt : décomposer dans le temps la réponse des ménages au choc sur les revenus.

$$z_{i,t} = \sum_{h=-10}^{+4} \beta_h \mathbf{1}_{\{e_{it}=h\}} + \delta_i + \gamma_t + \varepsilon_{i,t} \quad (4)$$

$$\Delta C_{i,t} = \beta (X_i) \Delta Y_{i,t} + \Gamma_t + \Delta \varepsilon_{i,t} \quad (5)$$

Les pertes d'emploi (2/2)



Résultats :

- MPC de 0.20 (écart-type : 0.02).
- Décroissante avec le patrimoine liquide.
- En moyenne, compensation de 16 % de la perte de revenus pré-chômage ($\simeq 1,6$ mois) par une baisse de la consommation ($\simeq 0,25$ mois).

Les hausses de salaires (1/2)

Contexte : Calculer la MPC dans le cas d'un choc positif et durable. Dans le cas américain, Baker (2018) étudie l'élasticité de la consommation au revenu à partir d'une hausse de revenus réguliers.

Concepts et champ : Nous définissons une hausse de salaire durable avec les conditions suivantes :

- le mois de la hausse $t = 0$ est défini par le mois avec la plus forte hausse relative du salaire
- monotonicité stricte entre les quartiles de revenus pré-hausse et post-hausse
- exclusion des faibles salaires (i.e $< 1\,000$ €)

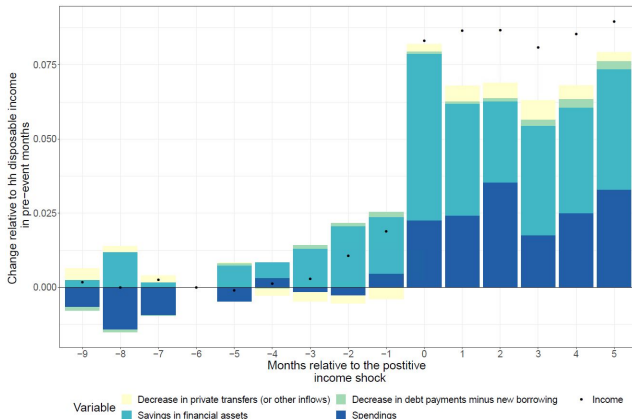
Stratégie d'identification : Appariement par score de propension entre les traités et des individus qui n'ont pas eu de hausses de salaires durables. $\simeq 3\,300$ traités et $\simeq 6\,400$ contrôles.

Estimation : Modélisation « event study » en différences de différences. Intérêt : décomposer dans le temps la réponse des ménages au choc sur les revenus.

$$z_{i,t} = \sum_{h=-9}^{+6} \beta_h T_i \mathbb{1}_{\{e_{it}=h\}} + T_i + \gamma_t + \varepsilon_{i,t} \quad (6)$$

$$\Delta C_{i,t} = \beta (X_i) \Delta Y_{i,t} + \Gamma_t + \Delta \varepsilon_{i,t} \quad (7)$$

Les hausses de salaires (2/2)



Résultats :

- MPC de 0.41 (écart-type : 0.04).
- Elle décroît avec le patrimoine liquide.
- En moyenne, hausse de revenus cumulée équivalente à 54 % du revenu mensuel pré-hausse + hausse de la consommation équivalente de 16 %.

Plan

Motivation

Forces et faiblesses des données bancaires

Résultats empiriques

Chocs transitoires

Chocs permanents

Conclusion

L'utilisation des données bancaires : un arbitrage entre atouts et fragilités ?

Sur les données bancaires :

- apport indéniable : haute fréquence, variété des études possibles (différents chocs, différents horizons temporels, différentes sous-population)
- mais aussi des fragilités (absence de labels, définitions des variables clés \neq définitions utilisées par le SSP, multibancarisation, représentativité)

Sur la problématique des MPC :

- peu de différences entre les chocs temporaires et permanents
- asymétrie selon le signe du choc de revenus (i.e $MPC_+ > MPC_-$) : interprétations comportementales (ex : aversion à la perte) possibles ?
- la MPC est toujours décroissante du niveau de patrimoine liquide : en lien avec les contraintes de liquidité ?

Merci !

Remerciements

Ce travail repose sur l'exploitation de données de comptes bancaires auxquelles le Crédit Mutuel Alliance Fédérale et La Banque Postale ont permis l'accès.

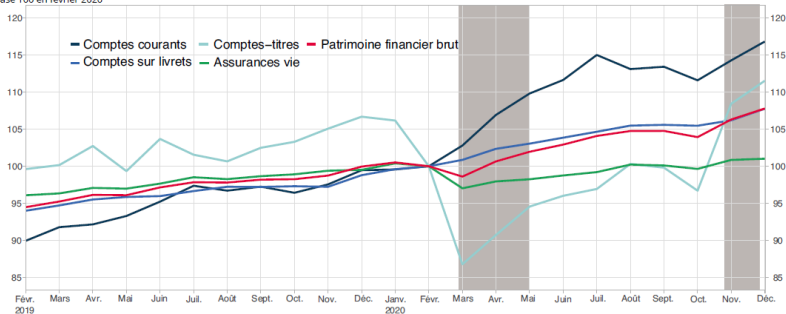
Nous remercions ces partenaires pour nous avoir facilité cet accès technique mais aussi pour leur disponibilité lors des échanges.

Annexe

Note de conjoncture de mars 2021

► 8. Évolution du patrimoine financier brut des ménages de l'échantillon redressé et de ses composantes en 2019 et en 2020

base 100 en février 2020



Lecture : en août 2020, le patrimoine financier brut était supérieur de 5 % au niveau de février 2020.

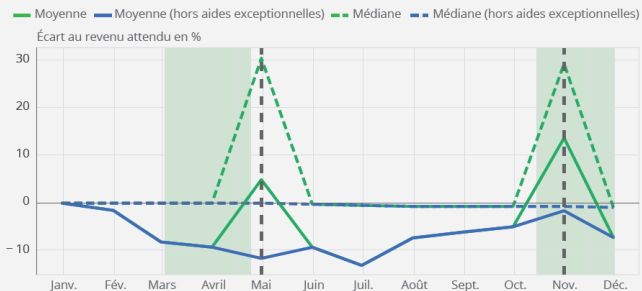
Note : les périodes de confinement sont indiquées par les zones grisées.

Source : données Crédit Mutuel Alliance Fédérale, calculs Insee

[retour](#)

Insee Analyses novembre 2021 sur la précarité

► 3. Écarts des revenus mensuels en 2020 par rapport au niveau simulé hors crise, pour certains allocataires du RSA en début d'année



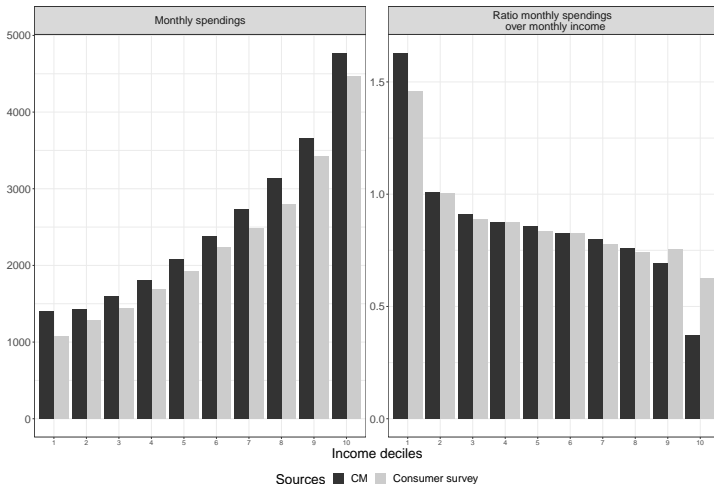
Notes : les revenus « hors aides exceptionnelles » excluent les aides exceptionnelles de mai et novembre 2020. Les périodes de confinements sont représentées par des bandes colorées.

Lecture : en avril 2020, les clients qui étaient seuls, sans personne à charge et percevaient le montant maximal du RSA en début d'année (janvier ou février) ont des revenus en moyenne 9 % inférieurs au niveau attendu si la tendance pré-crise s'était poursuivie.

Champ : France, échantillon de clients bancarisés principaux de *La Banque Postale* présents sur la période janvier 2019-juin 2021, seuls, sans personne à charge et percevant le montant maximal du RSA (avec ou sans forfait logement) en début d'année (en janvier ou en février).

Sources : La Banque Postale ; Insee : calculs des auteurs.

Dépenses de consommation : comparaison avec l'enquête Budget des Familles

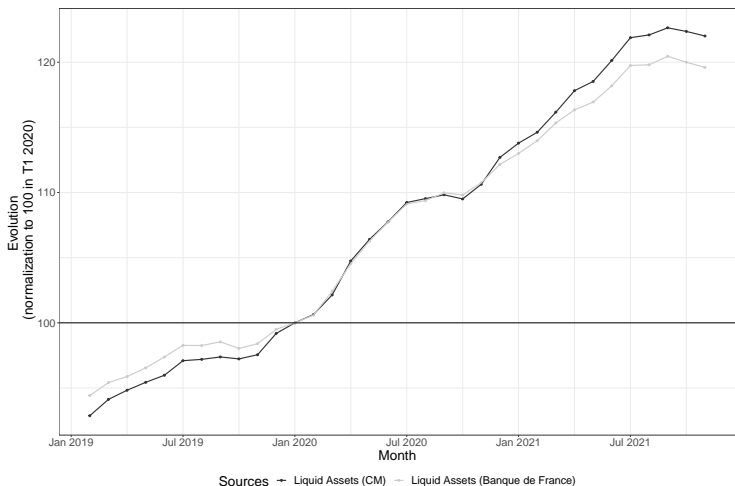


Statistiques descriptives sur l'échantillon du Crédit Mutuel

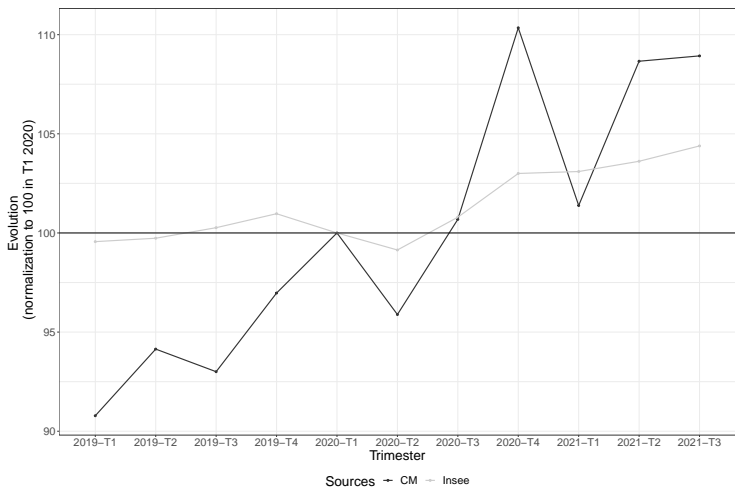
	(1)	(2)	(3)
	Crédit Mutuel		
	Unweighted sample	Weighted sample	National surveys
# of observations	169,163	169,163	
# of months	35	35	
	<i>Sample means</i>		
Spending	2,371	2,461	2,284
<i>Credit cards</i>	1,650	1,698	
<i>Bills</i>	713	756	
<i>Checks</i>	8	7	
<i>Utilities (bills and cards)</i>	140	148	113
<i>Groceries (cards)</i>	232	248	368
<i>Restaurants (cards)</i>	98	97	136
<i>Fuel (cards)</i>	78	79	92
Income	3,497	3,492	2,924
Financial Assets	50,657	55,615	50,882
<i>Liquid financial Assets</i>	32,858	35,241	24,270
Deposit account	9,514	10,525	4,046
Savings account	23,345	24,716	20,224
<i>Illiquid financial Assets</i>	17,799	20,374	26,612
Life insurance	13,597	15,748	18,947
Securities account	4,202	4,626	7,664
Monthly savings	95	83	
Loan net repayments	-390	-389	
Non-mortgage debt	-3,024	-3,086	-5,377
Mortgage debt	-34,793	-35,203	-38,605
Private transfers (or other inflows)	1,457	1,542	
Ratio liquid assets/deposit account	3.45	3.35	5.99
Age	49	52	52
Female	0.52	0.51	0.51
Craftsmen, merchants and business owners	0.06	0.06	0.04
Managerial and professional occupations	0.13	0.12	0.10
Technicians and associate professionals	0.14	0.13	0.13
Employees	0.24	0.22	0.14
Workers	0.13	0.12	0.11
Periphery areas	0.42	0.41	0.18
Rural areas	0.20	0.20	0.21
Urban areas	0.32	0.33	0.61

Pecuniary amounts : in €.

Patrimoine liquide brute agrégé : comparaison avec la Banque de France



Revenus agrégés : comparaison avec la comptabilité nationale



Effet du jour de paie : détails sur les concepts et le champ

Plusieurs filtres appliqués pour la construction du sous-échantillon de travail :

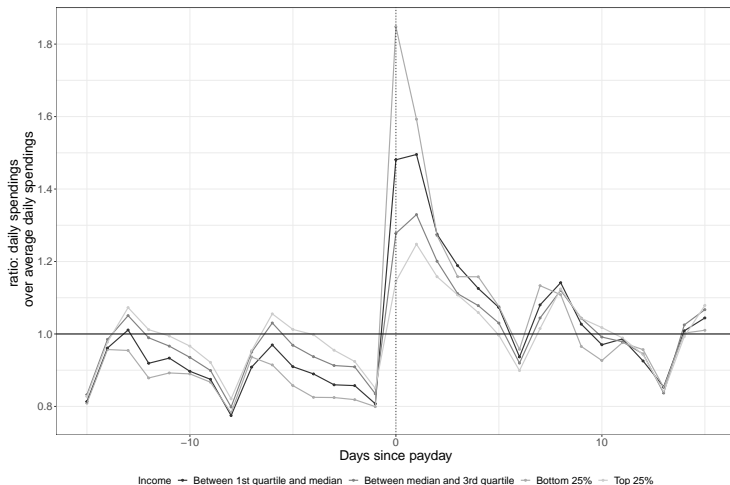
- restriction aux observations dont le revenu mensuel est versé un jour donné dans le mois.
- le revenu mensuel ne fait pas l'objet de plusieurs versements conséquents répartis sur l'ensemble du mois.
- les deux virements entrants les plus élevés sur le compte sont espacés d'au moins 25 jours et ne diffèrent pas de plus de 10 %.
- le montant de la paie doit être supérieur ou égal à 400 €.
- tous les critères précédents doivent être vérifiés au moins 3 mois consécutifs.

Au final, on obtient un sous-échantillon d'environ 81 000 clients différents \implies forte puissance statistique.

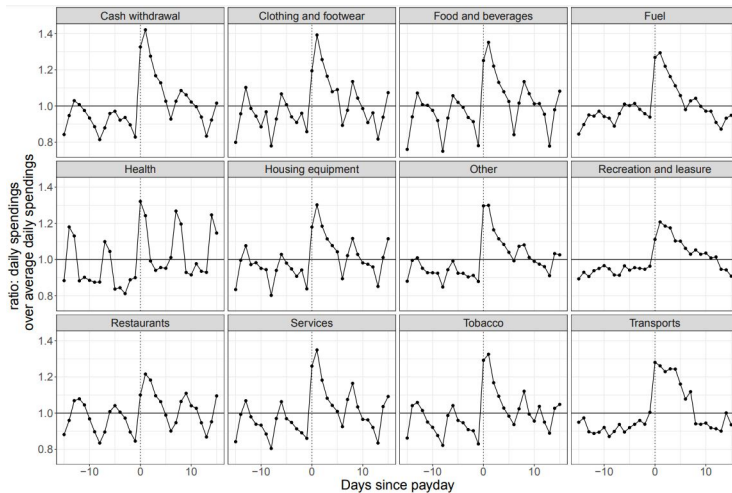
retour

Effet du jour de paie : détails des estimations

	All sample	Quartiles of income			
		1st	2nd	3rd	4th
Our estimates	0.43*** (0.01)	0.79*** (0.04)	0.46*** (0.02)	0.26*** (0.01)	0.14*** (0.01)
Olafsson-Pagel		0.88*** (0.01)	0.59*** (0.01)	0.44*** (0.01)	0.34*** (0.01)
Time FE	✓	✓	✓	✓	✓



Effet du jour de paie : résultats par catégories de biens



La perte d'emploi : détails sur les concepts et le champ

Définition d'une période de chômage :

- au moins deux mois de perception consécutive des allocations chômage.
- sans avoir perçu d'allocations chômage avant le début de la période (pendant au moins 10 mois).

Plusieurs filtres appliqués pour la construction du sous-échantillon de travail :

- exclusion des individus dont les revenus du travail mensuels ne dépassaient pas 1000 € au moins 5 mois sur les 10 mois précédant la période de chômage.
- sont également retirés les individus dont le revenu moyen entre $m = -10$ et $m = -4$ vaut plus de 95 % du revenu moyen entre $m = -1$ et $m = +1$.

retour

La perte d'emploi : analyse d'hétérogénéité

	(1)	(2)	(3)	(4)
	0.20*** (0.02)			0.31*** (0.04)
Financial assets				
<i>Bottom 25%</i>		0.33*** (0.03)		
<i>Between 1st quartile and median</i>		0.20*** (0.03)		
<i>Between median and 3rd quartile</i>		0.11*** (0.03)		
<i>Top 25%</i>		0.04 (0.04)		
Income				
<i>Bottom 25%</i>			0.25*** (0.04)	Ref.
<i>Between 1st quartile and median</i>			0.30*** (0.03)	0.04 (0.04)
<i>Between median and 3rd quartile</i>			0.20*** (0.03)	-0.02 (0.05)
<i>Top 25%</i>			0.08*** (0.03)	-0.09* (0.05)
Liquid financial assets				
<i>Bottom25%</i>				Ref.
<i>Between 1st quartile and median</i>				-0.09** (0.04)
<i>Between median and 3rd quartile</i>				-0.14*** (0.04)
<i>Top 25%</i>				-0.17*** (0.05)
Illiquid financial assets				
<i>No</i>				Ref.
<i>Yes</i>				-0.05 (0.04)
Consumption credit undertaken				
<i>No</i>				Ref.
<i>Yes</i>				0.02 (0.03)
Time FE	✓	✓	✓	✓
Socio-demographic controls	✓	✓	✓	✓
Financial pre-treatment vars	✓	✓	✓	✓
R^2	0.19	0.21	0.20	0.21
# of obs.	2,923	2,923	2,923	2,923

Les hausses de salaires : analyse d'hétérogénéité

	(1)	(2)	(3)	(4)
	0.41*** (0.04)			0.62*** (0.10)
Financial assets				
<i>Bottom 25%</i>		0.58*** (0.73)		
<i>Between 1st quartile and median</i>		0.47*** (0.07)		
<i>Between median and 3rd quartile</i>		0.36*** (0.07)		
<i>Top 25%</i>		0.18 (0.08)		
Income				
<i>Bottom 25%</i>			0.48*** (0.07)	Ref.
<i>Between 1st quartile and median</i>			0.49*** (0.08)	0.08 (0.11)
<i>Between median and 3rd quartile</i>			0.17* (0.09)	-0.20 (0.12)
<i>Top 25%</i>			0.43*** (0.11)	0.1* (0.14)
Liquid financial assets				
<i>Bottom 25%</i>				Ref.
<i>Between 1st quartile and median</i>				-0.21** (0.11)
<i>Between median and 3rd quartile</i>				-0.18*** (0.13)
<i>Top 25%</i>				-0.47*** (0.13)
Illiquid financial assets				
<i>No</i>				Ref.
<i>Yes</i>				-0.06 (0.10)
Consumption credit undertaken				
<i>No</i>				Ref.
<i>Yes</i>				0.01 (0.10)
Time FE	✓	✓	✓	✓
Socio-demographic controls	✓	✓	✓	✓
Financial pre-treatment vars	✓	✓	✓	✓
R^2	0.28	0.28	0.28	0.29
# of obs.	3,338	3,338	3,338	3,338