



HAL
open science

Essays in health and labor economics

Pauline Leveneur

► **To cite this version:**

Pauline Leveneur. Essays in health and labor economics. Economics and Finance. Institut Polytechnique de Paris, 2024. English. NNT : 2024IPPAG010 . tel-04904671

HAL Id: tel-04904671

<https://theses.hal.science/tel-04904671v1>

Submitted on 21 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT
POLYTECHNIQUE
DE PARIS

NNT : 2024IPPAG010

Thèse de doctorat



Essays in Health and Labor Economics

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École Nationale de la Statistique et de l'Administration Economique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat: Sciences Économiques

Thèse présentée et soutenue à Palaiseau, le 9 Décembre 2024, par

PAULINE LEVENEUR

Composition du Jury :

Sylvie Blasco Professeur d'Université, Université Caen Normandie	Présidente/Rapporteuse
Johan Vikström Professeur, IFAU et Uppsala University	Rapporteur
Mathilde Godard Chargée de recherche, CNRS & Université Paris-Dauphine-PSL	Examinatrice
Alexandra Roulet Professeur Adjoint, INSEAD	Examinatrice
Philippe Choné Professeur, CREST	Directeur de thèse
Arne Uhlendorff Directeur de recherche, CNRS & CREST	Co-directeur de thèse

Acknowledgments - Remerciements

Ce manuscrit est le résultat d'une belle aventure, non sans obstacle, mais ô combien enrichissante. Je profite de ce rare moment qui m'est offert pour remercier toutes les personnes rencontrées sur le chemin, ou déjà présentes, qui ont contribué à l'aboutissement de ce projet.

Mes premiers remerciements vont à mes directeurs de thèse, Philippe Choné et Arne Uhlendorff, pour leur accompagnement, leur soutien et leur écoute attentive tout au long des différentes étapes de la thèse.

Philippe, tu as accueilli avec enthousiasme ma volonté de revenir au CREST pour travailler sur des sujets en économie de la santé. Merci de m'avoir aidée à développer mon projet de thèse et de m'avoir permis de débiter dans de très bonnes conditions, notamment avec des données de qualité. Je te remercie également pour ton soutien, tes conseils et le partage de tes nombreuses connaissances sur les sujets santé. Ta rigueur scientifique et tes commentaires avisés ont été précieux pour faire progresser ma recherche.

Arne, thank you for your guidance in shaping my PhD journey and for mentoring me long before officially becoming my supervisor. Your advice and insightful feedback on my work have always been very accurate, and my research has greatly benefited from your constructive suggestions. I have learned a lot from our discussions on labor-related topics and empirical approaches. I am especially grateful for your constant availability, your support along the way, and your encouragement during moments of doubt.

I am deeply grateful to the members of the jury for accepting the invitation. Their work has been a source of inspiration, and I am very grateful to have had the opportunity to discuss

mine with them. Thank you to the reviewers, Sylvie Blasco and Johan Vikström, for providing insightful and constructive feedback to help me further develop these projects. Thank you also to Mathilde Godard and Alexandra Roulet for their helpful comments on my work.

Je voudrais chaleureusement remercier mes co-auteurs qui ont contribué à la réalisation de ces travaux. Merci à Olivier Allais, Céline Bonnet et Maxime Tranchard, avec qui j'ai commencé à travailler en tant qu'ingénieure d'études pour un projet sur l'obésité infantile ; je suis ravie que notre collaboration se soit poursuivie ensuite. Merci, Olivier, de m'avoir guidée dans l'exploration de ces thèmes de recherche et montré l'importance de mener des travaux ancrés dans les débats publics contemporains. Merci, Céline, de m'avoir formée aux méthodes d'estimation structurelle, ainsi que pour ta détermination inspirante, et tes qualités humaines qui m'ont permis d'évoluer en tant que chercheuse et en tant que personne. Merci, Maxime, pour ton enthousiasme constant à explorer de nouvelles idées et ta détermination à trouver des solutions pour les concrétiser. Je remercie aussi Pauline Lesterquy, avec qui j'explore le lien entre automatisation et santé, de m'avoir initiée aux mystères de la classification NC8, et pour qui les données comptables d'entreprises n'ont pas de secret.

Cette thèse a bénéficié de l'accès à divers jeux de données. Je remercie Marcel Goldberg et Marie Zins de m'avoir accordé l'accès à la cohorte épidémiologique Constances, ainsi que Julie Gourmelen et Sofiane Kab pour le partage de leur expertise sur les données et les échanges constructifs tout au long du projet, et Julie Rochut pour son aide avec les données de la CNAV. Ces travaux reposent également sur les données administratives d'entreprises de l'INSEE, du Ministère des Finances et de la Direction générale des Douanes et Drois indirects. Je remercie également l'INRAE de m'avoir permis d'utiliser les données d'achat de Kantar Worldpanel.

Ces trois années passées au CREST ont été exceptionnelles, et je tiens à remercier toutes les personnes qui contribuent à cet environnement remarquable ! Ma recherche a beaucoup bénéficié des réflexions de collègues brillants et des présentations inspirantes des invités. Au-delà de l'excellence académique, l'atmosphère chaleureuse, bi-

enveillante et encourageante du labo a été essentielle pour me permettre de me sentir constamment soutenue et pour prendre confiance.

Je voudrais tout d'abord remercier tous les professeurs du CREST. Plus particulièrement, je tiens à remercier les membres de mes comités de suivi, Roland Rathelot, Benoit Schmutz-Bloch et Sara Signorelli, pour leurs commentaires constructifs sur mes projets, pour leurs conseils sur l'organisation de cette thèse, ainsi que pour leur disponibilité. Je remercie également Grégory Corcos, Xavier D'Haultfoeuille, Bertrand Garbinti, Roxana Fernández, Florian Grosset-Touba, Yves Le Yaouanq, Marion Leroutier, Clément Malgouyres et Lionel Wilner d'avoir pris le temps de discuter de ma recherche.

J'ai beaucoup apprécié mon expérience d'enseignement à l'ENSAE, et je tiens à exprimer ma gratitude envers Laurent Davezies, Xavier D'Haultfoeuille et Michael Visser pour m'avoir donné l'opportunité d'être leur chargée de TDs. J'ai une pensée particulière pour mes étudiants ; la transmission de connaissances et nos échanges ont été véritablement enrichissants pour moi.

Je suis également profondément reconnaissante envers tous les doctorants que j'ai rencontrés, qui font partie intégrante de la communauté incroyable du labo. Être entourée de personnes aussi solidaires et inspirantes a joué un rôle majeur dans mes progrès. Merci de m'avoir aidée à réfléchir à mes projets et à affiner mes recherches à travers les feedbacks post-séminaires et les discussions informelles. Il en faut pour s'accrocher pendant ces semaines où rien ne semble marcher, mais les encouragements et le soutien constant que nous nous sommes apportés ont fait toute la différence. Un grand merci pour tous les moments joyeux que nous avons partagés, notamment les déjeuners au Magnan, les pauses ensoleillées sur la terrasse du 4ème étage, les trajets en RER retour vers Paris, les discussions quotidiennes à la machine à café, les afterworks, ainsi que les matchs de basket ou de pétanque ensemble !

J'ai eu la joie de partager cette expérience auprès de ma très chère amie Alice Lapeyre. Ta détermination et tes réflexions profondes n'ont cessé de m'inspirer. Merci d'être toujours à mes côtés, depuis de nombreuses années, et de m'aider à naviguer à travers beaucoup d'aspects de ma vie.

Je suis éternellement reconnaissante pour la *safe space* et le soutien sans faille ap-

portés par : Marion Brouard, qui a été présente dans toutes mes périodes difficiles, trouvant toujours les mots justes pour me remotiver, mais également dans les bons moments en célébrant et en me rappelant mes réussites — cette aventure aurait été bien différente sans toi ; Léa Dubreuil, dont l'énergie débordante et l'enthousiasme inspirant m'ont toujours boostée — je suis tellement heureuse d'avoir partagé mes aventures nordiques avec toi ; Clément Montes, qui illumine colore littéralement nos journées au labo — merci pour ton écoute et tes encouragements constants ; et Pedro Vergara Merino, dont la gentillesse et la présence bienveillante et attentionnée sont un véritable cadeau — sans oublier qu'il est un excellent partenaire de salsa.

J'ai également eu la chance de passer mes journées avec des collègues de bureau formidables. Un grand merci à Étienne Guigue pour m'avoir chaleureusement accueillie dans le bureau 4092, tes conseils m'ont beaucoup rassurée au début de la thèse ; nos discussions de travail, les débriefings hebdomadaires des matchs de basket et notre attachement commun au sud-ouest ont également été très appréciés ; *to Giulia Vattuone for your precious advice during my first year, it is always a pleasure to meet you again* ; à Elia Pérennes pour les moments agréables que nous avons partagés, et pour m'avoir gentiment cédé ton bureau ; à Arnaud Pandevant pour tes anecdotes insolites et pour avoir relevé le niveau en économie théorique du bureau ; à Mitia Oberti, autant connu pour ton énergie communicative que pour ta routine sportive impressionnante ; *to Jeremias Klæui for your good vibes and collective spirit, and for introducing us to Swiss specialties* ; à Louise Demoury pour ta joie, ton dynamisme quotidiens, et tes délicieuses recettes ; et à Yanis Zeghal pour t'assurer que nos estomacs ne soient jamais vides !

Je tiens également à remercier tous les doctorants que j'ai eu la chance de rencontrer au fur et à mesure des années, en particulier Maddalena Conte, Mario Herrera, Riccardo Norbiato et Felix Schleef *for the bonds we have built*, ainsi que Rémi Avignon, Chloé Antoine, Jeanne Astier, Brice Berland, Camille Boissel, Emma Bonutti, Léa Bou Sleiman, Wael Bousselmi, Guillaume Bied, Fanny Cartellier, Arnault Chatelain, Paul-Emmanuel Chouc, Pauline Carry, Gwen-Jiro Clochard, Héloïse Clolerly, Louis De Lachapelle, Julien Dubois, Sofian El Atifi, Andréa Epivent, Adélaïde Fabbi, Aurélien Frot, Morgane Guignard, Yannick Guyonvarch, Laure Heidmann, Morgane Hoffmann,

Myriam Kassoul, Margarita Kirneva, Raphaël Lafrogne-Joussier, Elisabeth Leduc, Titouan Le Calvé, Claire Leroy, Julia Nicolas, Elio Nimier-David, Denys Medee, Federica Meluzzi, Gaëtan Ménard, Matéo Moglia, Thomas Monnier, Inès Moutachaker, Martin Mugnier, Louis Pape, Félix Pasquier, Aisa Peng, Fabien Perez, Inès Picard, Pierre Rousseaux, Yuanzhe Tang, Jérôme Trinh, Johanne Trotin, Théo Valentin, Adrien Vallette, Vincent Verger, Dalil Youcefi et Yiyun Zheng. Je vous souhaite, à tous, tout le meilleur du plus profond de mon coeur !

J'aimerais également remercier toutes les personnes qui ont contribué à me permettre de réaliser ma thèse dans les meilleures conditions : Arnak Dalayan, Teddy Arrif, Adèle Ado, Pascale Deniau, Djamila Gherabi, Murielle Jules, Weronika Leduc, Eliane Madelaine, Delphine Maucherat, Leyla Marzuk, Alexandra Noiret, Philippe Pinczon du Sel, Lyza Racon, Fanda Traoré, Edith Verger, Arielle Vernet et Justine Viaud. Un merci tout particulier à Tristan Duchenne pour son écoute et son engagement dévoué au sein du laboratoire. Un grand merci également à Sylvie Coussot pour son accompagnement précieux.

I also had the fantastic opportunity to visit Uppsala University during the third year of my PhD. Thanks a lot to everyone I met at the Department of Economics, IFAU and the Centre for Health Economics Research. I am very grateful to Erik Grönqvist for hosting me and giving me constructive feedback on my projects. I also benefited a lot from interactions with Adrian Adermon, Jonas Cederlöf, Mikael Elinder, Peter Fredriksson, Georg Graetz, Lena Hensvik, Martin Nybom, Jan Sauermann and Johan Vikström.

I also want to thank the PhD students for welcoming me so warmly and introducing me to the fascinating life in Uppsala. I am especially grateful to Elisavet Iliadi for being such a wonderful and caring person, always looking out for me, and to Gilson Dutra, the sun of the department, for brightening our days and for his genuine generosity. You both made me feel at home. It was a real pleasure sharing my office with Sarah Gharbi, I truly enjoyed our time together and our conversations on many topics, including French data of course. A big thank you to Gustav Lundström for being very considerate and welcoming, and to William Skoglund for his kindness and contagious good mood, which made my stay even more enjoyable. Finally,

for all the lunches, fikas and drinks shared together, I would also like to thank Alexandra, Ricard, Martin, Olle, Jakob, Greta, Henning, Svante, Lovisa, Suna, Dogan, Oona, Emma, Thomas, Max, Levy, Andrii, Christian, Erik, Henrik, Xuemeng, Erika, Maria Elena, Caio, Alexander, Amin and Sébastien. Tusen tack!

I also want to thank all the people I had the pleasure of meeting during conferences, workshops, and seminars, who took the time to provide feedback on my projects and made me feel part of this community.

J'aimerais également remercier les personnes qui ont influencé mon parcours. Merci à Morgane Cure, Christophe Gaillac, Jérémy l'Hour, et Clémence Lenoir, dont les qualités d'enseignement m'ont particulièrement inspirée. Merci à Marine Guillerm et Mathilde Gaini pour leur encadrement lors de mon stage à la DARES et pour m'avoir transmis leur appétence pour l'évaluation de politiques publiques. *A big thank you to Bobby Mihaylova, who introduced me to the questions linking individual health and labor market outcomes during my internship at Queen Mary University back in 2017. This first research experience sparked my interest in this topic.* Merci également à l'équipe Food de la Toulouse School of Economics, Zohra Bouamra-Mechemache, Clara Etcheverry, Joel Jardi, Valérie Orozco, Olivier de Mouzon, Marine Spiteri et Vincent Réquillart, pour ces deux années très enrichissantes. Merci également à Martin Monziols, d'avoir partagé ses réflexions sur mon premier chapitre de thèse.

Je tiens également à remercier mes proches avec qui j'ai passé de très bons moments, loin des préoccupations sur mes projets de recherche.

Tout d'abord, un grand merci à mes amis pour les précieux moments partagés et pour m'avoir permis de me changer les idées. A mes amies de prépa : Cécile pour nos discussions et nos goûters parisiens et Laura pour nos appels habituels et pour les séjours bordelais bien appréciés; ainsi que Julien et Quentin que je prends toujours plaisir à retrouver. A Laure, pour les nombreux dîners et tous nos moments partagés. A Diane, Lucie et Myriam, devenues de véritables amies en dehors du terrain, sans oublier Fadel, Julien et Théo. Aux copains d'école, Coline, Grégoire, Alice S., Bruno,

Bertille, Elie, Romain et Loïc. A Camille, pour nos trajets partagés. A mon amie de longue date, Elodie. Aux fidèles albigeoises, Angélica et Léonie. A Thym et Pétro pour leur accueil toujours enthousiaste sur l'Ile intense. Au point d'ancrage toulousain formé par Eloïse, Constant, Lulu, Flo, Manon, Ben, Charlène, Alexis, Guéric, Myrtille et Loïs.

Je remercie également mes co-équipières de basket du PUC, Clarisse, Charlotte, Elsa, Emilie, Jeanne, Jen, Lucile, Lucille, Lucie D, Lucie G, Lydie, Marjorie, Milena, Muriel, Pauline, Sarah, Raphaëlle et Rachel. Merci aussi aux coachs Jérémy, Louis, Anton et Fodie pour leur confiance et pour m'avoir fait progresser. Les entraînements, matchs ainsi que multiples verres et soirées ont été une vraie bulle d'air et un bon moyen de décompresser pendant cette période. Merci également à mes anciennes coéquipières du BP14 et les copains de street à Cité U. *Finally, a huge thanks to the Campus 1477 team, especially Anne, Kendra, and William for making me feel part of the group; the pick-up sessions warmed my evenings in Uppsala.*

Enfin, je tiens à profiter de ce moment pour remercier ma famille d'avoir été présente pour moi et de m'avoir tant apporté. Merci à mes parents, Thérésien et Anne, de m'avoir toujours encouragée dans mes choix. Papa, merci d'être toujours si fier de tout ce que j'entreprends et de m'avoir montré ce qui rend heureux. Maman, merci d'être toujours là pour moi ; j'espère un jour pouvoir refléter ta détermination et ton courage. Merci à Arnaud, mon petit frère - mais grand d'esprit et de coeur - avec qui je peux tout partager et qui n'oublie jamais de me challenger pour m'aider à avancer. Merci à mes grands-parents, Pierre et Josette, de nous avoir transmis tout cet amour qui vit en moi pour toujours. Merci à ma belle-soeur, Laurie, pour tous les bons moments partagés ensemble. Merci à mes beaux-parents, Thierry et Sylvie, de m'avoir accueillie et de me soutenir. Les nombreux séjours en terres tarnaises ont toujours été l'occasion de déconnecter et de me ressourcer.

Robin, merci pour ton soutien précieux depuis tant d'années et de toujours croire en moi, bien au-delà de la thèse. Par ta bienveillance, tes encouragements et ton amour, tu m'as donné de la force et aidée à prendre du recul pendant cette période. Surtout, merci de contribuer à mon bonheur, en me faisant rire au quotidien, et en rendant chaque journée plus belle et plus légère.

Table of Contents

Acknowledgments - Remerciements	3
Summary	17
Résumé	21
General Introduction	25
1 Measuring Individual Health	26
1.1 Objective and subjective health	26
1.2 Mental and physical health	26
1.3 Health status and healthcare use	27
1.4 Some figures in France	27
2 The Interplay between Work and Health	28
2.1 The effect of employment on health	28
2.2 The career shock of job loss	29
2.3 The changing nature of work: technological advancements and their impacts	29
2.3.1 A brief history of technological change	30
2.3.2 How will the workforce be affected?	30
2.3.3 Should workers worry about automation and digital technologies?	32
2.3.4 Should we expect effects of automation on workers' health?	33
3 Tackling the Obesity Epidemic: the Case of Sugar Taxes	34
3.1 Trends in obesity and policy tools	34

3.2	The challenges of nutritional taxes	36
4	Methodology of the Thesis	38
4.1	Data sources	38
4.2	Empirical methods	39
5	Summary of the Chapters and Contributions	40
5.1	Chapter 1 - Job Loss and Health Outcomes: Evidence from France	40
5.2	Chapter 2 - Automation at the Workplace: Implications for Workers' Health	42
5.3	Chapter 3 - Are Sugar Taxes Well Designed? Empirical evidence from the UK soft drink market	44
	Introduction générale	47
	1 Job Loss and Health Outcomes: Evidence from France	79
1	Introduction	81
2	Institutional background and data	84
2.1	Institutional background	84
2.2	Data	85
3	Empirical strategy	93
3.1	Main specification	93
3.2	Alternative specifications and robustness checks	95
4	Results	97
4.1	Effect of job loss on Economic Outcomes	97
4.2	Effect of job loss on Mental Health Outcomes	98
4.3	Effect of job loss on other Health Outcomes	101
5	Channels	103
6	Conclusion	105
	Appendices to Chapter 1	107
	2 Automation at the Workplace: Implications for Workers' Health	127
1	Introduction	129
2	Data	133

2.1	Worker-level data	133
2.2	Firm-level data	135
2.3	Proxies for automation in the firm	136
2.3.1	Stock of industrial equipment and machines	136
2.3.2	Imports of automation technologies and industrial robots	137
2.3.3	Aggregate trends and comparison of the proxies	139
3	Stylized facts and treatment definition	140
3.1	Stylized facts	140
3.2	Treatment definition: characterizing a technological shift	143
3.3	Samples of firms and workers	145
3.3.1	Firms	145
3.3.2	Workers	145
3.4	Descriptive statistics	147
4	Firm-level results	151
4.1	Specification	151
4.2	Results	153
4.2.1	Validation exercise	153
4.2.2	Main results	154
4.3	Discussion	156
4.3.1	Robustness checks	156
4.3.2	Changing automation event	156
4.3.3	Changing specification	157
5	Worker-level analysis	157
5.1	Specification	158
5.2	Results on Labor market outcomes	159
5.3	Results on Healthcare outcomes	160
5.3.1	Mental health	161
5.3.2	Physical health	161
5.3.3	Health insurance	163
5.4	Discussion	164

6	Further steps	165
6.1	Improving the identification strategy to estimate causal effects . .	165
6.1.1	Find an instrumental approach	165
6.1.2	Refine the worker-level identification and analysis . . .	167
6.2	Decomposing between different types of automation technologies	168
7	Conclusion	168
	Appendices to Chapter 2	171

3	Are Sugar Taxes Well Designed? Empirical evidence from the UK soft drink market	211
1	Introduction	213
2	The UK soft drink market	216
2.1	Data	216
2.2	Descriptive statistics	217
3	Methodology	218
3.1	Welfare maximization program	218
3.1.1	Definition of a tax design	218
3.1.2	Theoretical welfare definition	219
3.2	Consumer and firms behaviors	220
3.2.1	Demand model	220
3.2.2	Supply model	223
3.2.3	Results of the model	224
3.3	Counterfactual simulations	225
3.3.1	Objective function	226
3.3.2	New price equilibrium	227
4	Results	228
4.1	Optimal design of sugar taxes	228
4.2	Validity of our methodology	231
4.2.1	Comparison of strategic vs. non-strategic tax designs . .	231
4.2.2	Discussion on the value of the social cost of sugar	233
4.3	Comparison with the SDIL	233

5	Conclusion	235
	Appendices to Chapter 3	237
	Bibliography	261
	List of Figures	265
	List of Tables	268

Summary

This thesis explores key determinants of individual health, focusing on two main areas: the consequences of career shocks on workers' health and the design of nutritional fiscal measures. The first objective is to deepen our understanding of how employment status and conditions affect health, while the second is to assess the welfare effects of fiscal policies aimed at reducing obesity.

The first chapter investigates the short- and long-term health impacts of job loss in France from 2007 to 2019. Job loss often leads to income reduction, economic stress, and social isolation. Although unemployment benefits help cover income loss, the health risks associated with job displacement are frequently overlooked. The analysis leverages exogenous job losses identified through plant closures. Using individual healthcare data from the *Constances* epidemiological cohort, this study provides causal estimates by comparing the health outcomes of displaced workers with those of non-displaced workers through a difference-in-differences setting. The findings reveal that displaced workers experience long-lasting declines in both earnings and health. Mental health suffers notably, with persistent increased use of antidepressants, anxiolytics and sleeping pills. Physical health shows short-term deterioration, evidenced by higher rates of visits to doctors and hospitalizations. Vulnerable groups, such as older and low-skilled workers, are particularly affected. Among the mechanisms, the prospect of re-employment significantly mitigates the detrimental impact on mental health.

The second chapter examines the health impacts of workplace automation and robotics in the French manufacturing sector between 2007 and 2019. Automation's effects on workers' health are mixed, with potential risks stemming from job displacement and

increased work intensity, but also benefits such as reduced physical strain and improved safety. Using firm-level administrative data, this study tracks automation investments through proxies such as the stock of industrial equipment and imports of automation goods. The results suggest that productivity gains from automation outweigh the displacement effects at the firm-level, leading to an increase in employment, particularly among blue-collar workers. By linking firm-level automation to worker-level labor and health outcomes, the study compares workers of the *Constances* cohort highly exposed to a workplace automation shock with those less exposed, using a difference-in-differences framework. We show that automation does not significantly affect incumbent workers' earnings or employment status. However, it is associated with a rise in the use of antidepressants and anxiolytics. No significant changes are observed in physical health outcomes on average. Moreover, we document that older workers are adversely affected by workplace automation.

The third chapter examines the optimal design of sugar taxes in the UK soft drink market. While economic theory advocates for a linear tax structure, many countries have implemented tiered designs (*i.e.*, tax rates that increase progressively based on sugar content). This study introduces a regulatory framework to evaluate the welfare effects of different tax structures, considering consumer surplus, firm profits, government revenue, and the externalities associated with excessive sugar consumption. Using home-scan consumer data and a structural discrete choice model, we simulate the impact of various tax scenarios on overall welfare, accounting for consumer preferences and firms' price responses. The findings reveal that tiered taxes, when integrating firms' strategic responses, can yield significantly greater welfare gains than both the UK's currently implemented Soft Drinks Industry Levy and a linear tax structure. The optimal design imposes higher taxes on high-sugar products, incentivizing firms to reduce the prices of lower-sugar alternatives.

This thesis adopts an empirical approach, drawing on a diverse set of data (including firm- and individual-level administrative records, survey data, and household purchase data) and leveraging reduced-form and structural estimations, to examine public health policies and labor market shocks.

Résumé

Cette thèse explore des déterminants fondamentaux de la santé individuelle en se concentrant sur deux axes principaux : les conséquences des chocs de carrière sur la santé des travailleurs et la conception de mesures fiscales nutritionnelles. Le premier objectif est d’approfondir notre compréhension de l’impact du statut et des conditions d’emploi sur la santé, et le second vise à évaluer les effets des politiques fiscales visant à réduire l’obésité sur le bien-être global.

Le premier chapitre examine les impacts de la perte d’emploi sur la santé à court et long terme en France entre 2007 et 2019. La perte d’emploi entraîne souvent une réduction de revenus, un stress économique et un isolement social. Bien que les allocations chômage contribuent à compenser la perte de revenu, les risques pour la santé associés au licenciement sont fréquemment négligés. L’analyse exploite les pertes d’emploi exogènes identifiées à travers les fermetures d’établissements. En utilisant des données individuelles de soins de santé issues de la cohorte épidémiologique *Constances*, cette étude fournit des estimations causales en comparant les indicateurs de santé des travailleurs licenciés à ceux des travailleurs non licenciés dans un cadre de différence-en-différences. Les résultats révèlent que les travailleurs licenciés subissent une baisse durable de leurs revenus et de leur état de santé. La santé mentale est particulièrement affectée, avec une augmentation persistante de l’utilisation d’antidépresseurs, d’anxiolytiques et de somnifères. La santé physique montre une détérioration à court terme, comme en témoigne une hausse des consultations médicales et des hospitalisations. Les groupes vulnérables, tels que les travailleurs âgés et peu qualifiés, sont particulièrement touchés. Parmi les mécanismes, la perspective de retrouver un emploi atténue significativement l’impact négatif sur la santé mentale.

Le deuxième chapitre examine les impacts de l'automatisation et de la robotisation au travail sur la santé des travailleurs dans le secteur manufacturier français entre 2007 et 2019. Les effets de l'automatisation sur la santé des travailleurs sont variés, avec des risques potentiels liés au licenciement et à l'intensification du travail, mais aussi des avantages tels que la réduction de la pénibilité physique et l'amélioration de la sécurité. En utilisant des données administratives au niveau des entreprises, cette étude suit les investissements en automatisation avec des indicateurs comme le stock d'équipements industriels et les importations de biens d'automatisation. Les résultats suggèrent que les gains de productivité liés à l'automatisation compensent les effets de destruction d'emplois au niveau de l'entreprise, conduisant à une augmentation de l'emploi, notamment parmi les ouvriers. En reliant l'automatisation au niveau des entreprises aux carrières et aux consommations de soins des travailleurs, l'étude compare des travailleurs de la cohorte *Constances* très exposés à un choc d'automatisation sur le lieu de travail à ceux peu exposés dans un cadre de différence-en-différences. Nous montrons que l'automatisation n'affecte pas significativement les revenus ou le statut d'emploi des travailleurs en poste avant l'évènement. Cependant, elle est associée à une augmentation de l'utilisation d'antidépresseurs et d'anxiolytiques. Aucun changement significatif n'est observé en moyenne pour les indicateurs de santé physique. De plus, nous mettons en évidence que les travailleurs plus âgés sont négativement affectés par l'automatisation sur le lieu de travail.

Le troisième chapitre examine la conception optimale des taxes sur le sucre dans le marché des boissons rafraîchissantes sans alcool au Royaume-Uni. Alors que la théorie économique préconise une structure de taxe linéaire, de nombreux pays ont mis en place des systèmes par paliers (c'est-à-dire des taux de taxe qui augmentent progressivement en fonction de la teneur en sucre). Cette étude introduit un cadre réglementaire pour évaluer les effets sur le bien-être de différentes structures de taxe, en prenant en compte le surplus des consommateurs, les profits des entreprises, les recettes de l'État, ainsi que les externalités liées à une consommation excessive de sucre. En utilisant des données de consommation à domicile et un modèle structurel de choix discret, nous simulons l'impact de divers scénarios de taxe sur le bien-être global, en ten-

ant compte des préférences des consommateurs et des réponses en prix des entreprises. Les résultats révèlent que les taxes par paliers, en intégrant les réponses stratégiques des entreprises, peuvent générer des gains de bien-être nettement supérieurs à ceux de la Soft Drinks Industry Levy actuellement appliquée au Royaume-Uni et d'une structure de taxe linéaire. La conception optimale impose des taxes plus élevées sur les produits à forte teneur en sucre, incitant ainsi les entreprises à réduire les prix des alternatives moins sucrées.

Cette thèse adopte une approche empirique en s'appuyant sur un ensemble varié de données (comprenant des données administratives au niveau des entreprises et des individus, des données d'enquête et des données sur les achats des ménages), et en mobilisant des estimations en forme réduite et structurelle, afin d'examiner les politiques de santé publique et les chocs sur le marché du travail.

General introduction

Individuals' health is shaped by a wide range of interconnected factors, ranging from biological elements to broader social and environmental conditions. Some of these determinants remain unexplained or beyond one's control, such as genetic predispositions and complex social dynamics. However, other causes have been well identified, including working conditions, income levels, access to healthcare, and lifestyle factors such as diet and physical activity, and their effects have been extensively studied. Economists play a crucial role in this context: by analyzing the predictable and measurable causes of health issues, they can propose public policies aimed at improving population well-being, particularly by addressing inequalities and mitigating risks associated with specific environments.

This dissertation lies at the intersection of health economics with two areas: labor markets and industrial organization. The goal is to deepen our understanding of two key factors that influence both individual and public health. First, I explore how certain career events, such as job loss (Chapter 1) and workplace automation (Chapter 2), impact both physical and mental health of workers. Second, I examine the design of sugar taxes and its implications for overall welfare (Chapter 3).

This introduction is structured into five sections. The first section discusses the challenges associated with measuring individual health. The second section explores the impact of the labor market on workers' health. The third section describes the obesity epidemic and focuses on one policy response: nutritional taxes. The fourth section provides an overview of the data sources and empirical methods used. Finally, the fifth section summarizes each chapter, highlighting their contributions to the existing literature.

1 Measuring Individual Health

The World Health Organization (WHO) defines health as "a state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity". This definition underscores that health is a complex construct, shaped by a combination of experiences and perceptions. There are several ways to measure health outcomes in data, each method offering its own strengths and limitations.

1.1 Objective and subjective health

Objective health refers to quantifiable metrics such as clinical indicators (*e.g.*, blood pressure, cholesterol levels, body mass index), disease diagnoses (*e.g.*, diabetes, heart conditions), or health outcomes like mortality and hospitalization rates. These measures offer a clear, standardized perspective on health, typically obtained from administrative data (*e.g.*, doctor visits, hospitalizations, prescribed drugs), medical diagnoses, biomarkers, or physical assessments.

In contrast, subjective health is based on individuals' perceptions and self-reports regarding their well-being. This data is usually collected through surveys or questionnaires where individuals evaluate their own health, mental state, or daily functioning. Common tools include self-rated health questions ("How would you rate your overall health?") and scales assessing mental well-being (*e.g.*, levels of stress, anxiety, or life satisfaction). Although subjective measures may lack the precision of clinical data, they provide valuable insights into how individuals perceive their health, encompassing broader impacts on quality of life. Importantly, subjective health data allows for the study of a different selection of individuals, as it does not solely rely on individuals visiting healthcare providers.

1.2 Mental and physical health

Physical health is generally assessed through objective indicators such as blood pressure, body mass index (BMI), cholesterol levels, or the presence of chronic diseases like diabetes or cardiovascular conditions. These indicators are commonly derived from medical records, health examinations, or administrative data.

On the other hand, mental health is more challenging to measure objectively and is often evaluated through self-reported surveys or psychological assessments (*e.g.*, standardized scales like the GHQ-12 or PHQ-9) that address symptoms of depression, anxiety, or overall emotional well-being. However, healthcare use, such as prescriptions for antidepressants, can provide additional insights into current or past mental health issues.

Despite being measured separately, physical and mental health are closely interconnected, as mental health issues can affect physical health outcomes and chronic physical conditions can significantly impact mental well-being.

1.3 Health status and healthcare use

Healthcare utilization refers to the use of medical services, often recorded in administrative data. While healthcare consumption can reflect health needs and is frequently used as a proxy for health, it does not always accurately capture an individual's health status. For example, high consumption of antidepressants might suggest improved mental health due to effective treatment; but it could also indicate ongoing mental health challenges requiring continual medication. Conversely, low consumption of antidepressants does not necessarily mean good mental health, as barriers like access to care, stigma, or financial constraints might prevent individuals from seeking necessary treatment.

In this dissertation, I rely mostly on objective health measures provided by reimbursement data, which offer detailed records of visits to doctors, hospital stays, dispensed drugs and sick leaves for both physical and mental health.

1.4 Some figures in France

The evolution of chronic diseases has steadily increased over the past few decades in France. In 2021, approximately 30% of the French adult population was living with at least one chronic disease, cardiovascular diseases, diabetes, and respiratory disorders being the most prevalent [[Santé Publique France, 2021](#)]. Approximately 20% of the French population experiences mental health disorders at some point in their lives, with anxiety and depression being the most common [[INSERM, 2020](#)]. Health expen-

ditures in France have risen consistently over the past decades, reaching nearly 12% of GDP in 2022. This makes France one of the countries with the highest healthcare spending in Europe. The rise in costs is largely attributed to an aging population, expensive medical innovations, and growing demand for chronic care. Most of the expenditures are covered by the national health insurance system, with a growing share directed towards outpatient care, pharmaceuticals, and hospital expenses [[European Observatory on Health Systems and Policies, 2023](#)].

2 The Interplay between Work and Health

2.1 The effect of employment on health

Employment and health are intrinsically linked throughout life. Individuals facing health issues often find it challenging to secure stable employment. Conversely, the nature of employment and working conditions can significantly affect health in various, sometimes ambiguous, ways.

Employment can have significant positive consequences on health. Beyond serving as a source of income, it provides individuals with a sense of purpose, structure, and social interaction, all of which contribute to better mental and emotional well-being. Furthermore, work-related activities often encourage physical activity and stimulate cognitive engagement. Additionally, in some contexts, jobs offer access to healthcare benefits, such as insurance coverage that includes preventive screenings and treatments, which can enhance physical health [[Burgard and Lin, 2013](#); [Marmot and Stansfeld, 2003](#)].

However, certain working conditions can negatively impact health. High levels of stress, long working hours, and job insecurity are associated with poorer mental health outcomes, including anxiety, depression, and burnout [[Berniell and Bietenbeck, 2020](#); [Blasco et al., 2022](#); [Caroli and Godard, 2016](#)]. Additionally, physically demanding or unsafe working environments can result in injury or chronic health issues. Sedentary lifestyles, which are common in office-based jobs, and the struggle to maintain a healthy work-life balance can further exacerbate health problems, contributing to con-

ditions such as cardiovascular disease, musculoskeletal disorders, and obesity.

I aim to shed light on the impact of two career shocks on workers' mental and physical health. In Chapter 1, I focus on involuntary job loss, which primarily represents a shock along the extensive margin. In Chapter 2, I investigate automation-driven changes in working conditions, along with job displacement or reallocation.

2.2 The career shock of job loss

One common and relatively short-term career shock is job loss, which can have immediate and long-lasting effects on individuals. This shock plays on the extensive margin of employment. In France, approximately 2.9 million people were unemployed in 2023, representing about 7.2% of the active population. While some workers may recover quickly, the consequences of job loss can persist, affecting both income and overall well-being. Job loss has significant economic consequences, notably through reductions in income, which can lead to long-term financial instability. Displaced workers experience persistent earnings losses, with incomes often remaining lower even several years after job displacement [[Jacobson et al., 1993](#)]. Many workers face difficulties in recovering pre-displacement wages, leading to heightened economic vulnerability.

Beyond income, job loss triggers major lifestyle changes, including increased stress, uncertainty, and shifts in household dynamics, which can result in reduced social engagement and changes in consumption patterns. These disruptions may contribute to a decline in mental and physical health, as financial strain and stress have been shown to increase the risk of depression, anxiety, and even chronic conditions like heart disease [[Sullivan and von Wachter, 2009](#)].

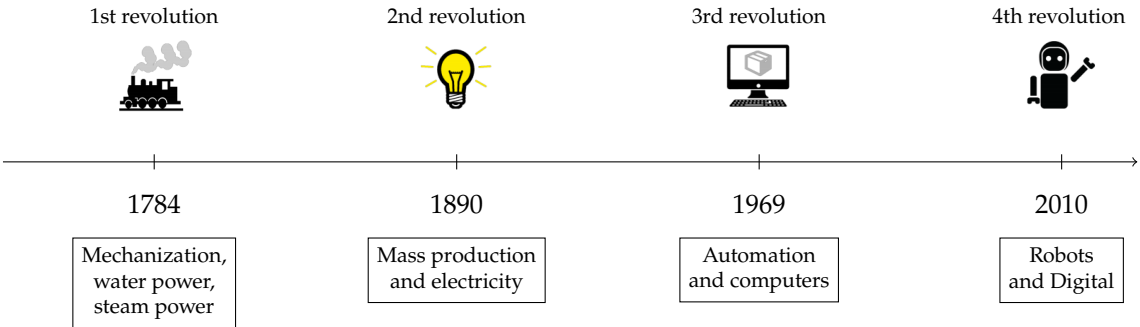
2.3 The changing nature of work: technological advancements and their impacts

The labor market has experienced significant structural changes in recent years, largely driven by technological innovation. These advancements have not only redefined the nature of work but also altered the composition and complexity of tasks.

2.3.1 A brief history of technological change

The timeline of industrial revolutions depicted in Figure 1 illustrates four major phases of technological and industrial transformation that have shaped modern economies. The first industrial revolution ("steam engine revolution"), beginning in 1784, was characterized by the introduction of mechanization, the use of water and steam power, and the advent of the first factories. This laid the groundwork for mass production and the widespread use of electricity during the second industrial revolution ("electricity revolution") in the late 19th century. In 1969, the third industrial revolution ("automation revolution") emerged with the rise of automation and the advent of computers, fundamentally changing the nature of work and production. Finally, the fourth industrial revolution ("digitalization revolution"), starting in 2010, is marked by advancements in robotics and digital technologies, which continue to push the boundaries of automation and connectivity in the global economy. In this dissertation, I focus on the two most recent industrial revolutions and the adoption of automation, robotics and digitalization at the workplace.

Figure 1: Timeline of industrial revolutions

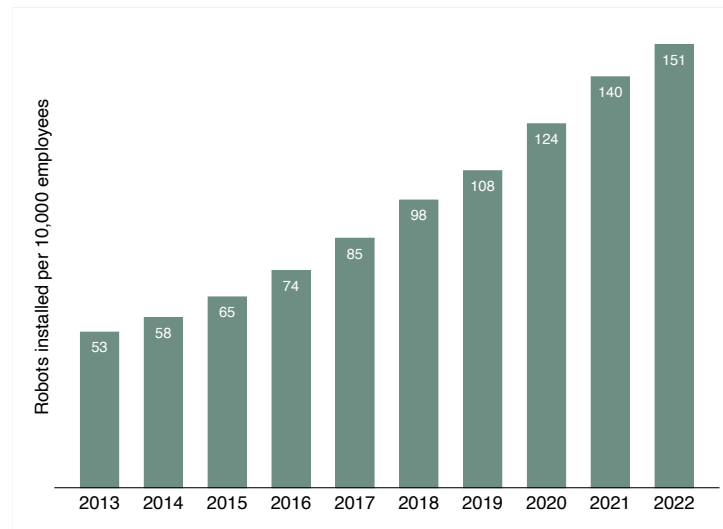


2.3.2 How will the workforce be affected?

In recent years, the penetration of industrial robots in companies has risen significantly in the world. Robot density, which stood at 53 robots per 10,000 employees in 2013, surged to 151 by 2022 (Figure 2). This sharp increase underscores the growing reliance on automation across various industries and suggests that workers are increasingly exposed to its presence in the workplace. The adoption of these innovations at the workplace is likely to affect a significant portion of the workforce. Almost half of all

jobs are at risk of being automated in the United States [Brynjolfsson and McAfee, 2014; Frey and Osborne, 2017] and up to 60% of jobs in OECD countries [World Bank, 2016]. A recent survey in France indicates that more than one third of employees work at a pace monitored by computers in France [DARES, 2021].

Figure 2: Trends in robot density



Notes: This figure plot trends in robot density from 2013 to 2022 in the World. Robot density is computed as the number of installed industrial robots per 10,000 employees. Source: World Robotics 2023, International Federation of Robotics.

The influential work of Autor et al. [2003] provides valuable insights into how these innovations may reshape the nature of work and alter the demand for various types of labor. They classify tasks along two key dimensions: cognitive vs. manual, and routine vs. non-routine. Cognitive tasks require problem-solving and complex decision-making, while manual tasks involve physical work. Routine tasks are repetitive and follow a set of predefined steps, whereas non-routine tasks are more complex and require flexibility, problem-solving and adaptability. By their nature, routine tasks are more likely to be subject to automation. For example, computers are more likely to substitute for workers performing routine, codifiable tasks, while serving as a complement to and enhancing workers' capabilities for non-routine cognitive tasks (*e.g.*, problem-solving and creativity) and manual tasks (*e.g.*, requiring situational adaptability) [Autor et al., 2003]. The task model developed in Acemoglu and Autor [2011] rationalizes this routine-biased technological change and how workers are differently affected depending on their initial skills.

This decomposition helps to understand the phenomenon of job polarization [Au-

tor and Dorn, 2013; Goldin and Katz, 1998; Goos, 2018; Goos and Manning, 2007]. In recent decades, a notable trend has been the polarization of the labor market, where wage increases have disproportionately favored workers at both the top and bottom of the income and skill distribution, leaving those in the middle behind. Technological advancements can be one reason for this phenomenon. For example, digital capital has a comparative advantage in performing routine-intensive tasks that can be easily codified in software, leading to the displacement of medium-skilled workers from these intermediate jobs. As a result, some medium-skilled workers shift their labor to simpler tasks, while others move into more complex roles.

2.3.3 Should workers worry about automation and digital technologies?

The rapid spread of automation and new technologies has stirred deep concerns among workers, raising fears of human labor becoming redundant. Following the "electricity revolution", Keynes wrote in 1930 *"We are being afflicted with a new disease of which some readers may not yet have heard the name, but of which they will hear a great deal in the years to come – namely, technological unemployment"*. In 1961, TIME Magazine published an article titled "The Automation Jobless," warning about the potential for mass unemployment due to technological advancements [Autor, 2015]. These fears of widespread displacement arise when automation operates at the extensive margin, creating substitution effects as technologies take over tasks previously performed by workers at a cheaper cost. This, in turn, leads to job destruction and a decrease in labor demand, ultimately putting downward pressure on wages [Acemoglu and Restrepo, 2020; Bessen et al., 2023; Graetz and Michaels, 2018]. For instance, consider the automotive manufacturing sector, where robotic systems have largely replaced manual labor in tasks like welding, painting, and assembly. Once performed by human workers, these tasks are now automated, reducing the need for a large workforce and pushing wages down for remaining labor-intensive roles.

Yet, history shows that human labor has not become obsolete [Autor, 2015]. Despite significant advancements in automating tasks, aggregate employment has not declined. Several mechanisms counterbalance the displacement effect. One key expla-

nation is that automation can complement human labor. The productivity effect plays a crucial role here: as automation lowers production costs, it increases demand for non-automated tasks [Acemoglu et al., 2020; Aghion et al., 2024; Bonfiglioli et al., 2024]. Additionally, automation often operates at the intensive margin, enhancing existing tasks rather than simply replacing human labor. Finally, the prevalent mechanism is the creation of new tasks, where human labor retains a comparative advantage over machines, fostering new employment opportunities [Acemoglu and Restrepo, 2019]. In the automotive industry, for instance, while robots have taken over repetitive tasks like welding or painting, human workers are still needed for complex decision-making and quality control, such as in design, engineering and inspections, where their judgment and skills are critical.

These examples illustrate that even when certain tasks are automated, workers may still be required to carry out complementary tasks. Instead of becoming obsolete, workers may need to adapt by applying different skills to the evolving tasks within their roles. As jobs transform, labor reallocation will occur: some workers will experience changes in their current occupations, while others may discover that different roles are now a better fit for their skills. The transition can be slow and painful, particularly for workers whose skills do not match the demands of the new technological landscape [Acemoglu and Restrepo, 2020; Bessen et al., 2023; Bonfiglioli et al., 2024; Humlum, 2021; Yakymovych, 2022]. This mismatch between technology and skills is a major challenge, as workers displaced by automation may not immediately possess the skills needed for the newly created jobs.

2.3.4 Should we expect effects of automation on workers' health?

The changes in tasks and working conditions may impact workers' physical and mental health, with effects that could vary based on individual worker characteristics. These shifts are expected to influence the prevalence of physical tasks, the frequency of workplace injuries, and overall work intensity. The sense of meaning at work might be diminished by increased automated monitoring, or it could improve if automation allows workers to perform more fulfilling tasks. Additionally, stress-related concerns

may arise from heightened production targets or anxiety over job security, as workers perceive competition with robots and automated systems.

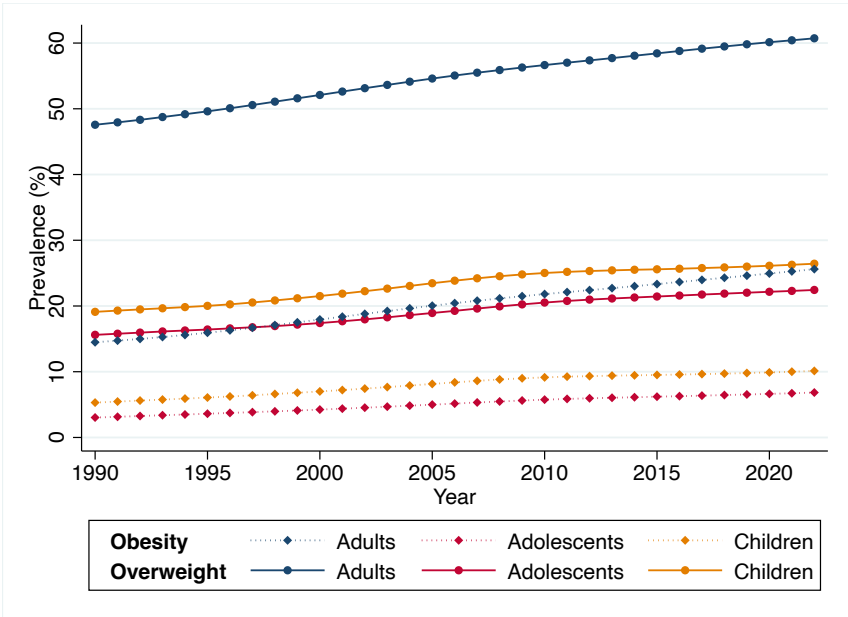
In Chapter 2, I quantify the effects of workplace automation and robotics on workers' health and investigate the mechanisms.

3 Tackling the Obesity Epidemic: the Case of Sugar Taxes

3.1 Trends in obesity and policy tools

Recent years have seen an alarming increase in obesity and overweight rates across all age groups in many countries. The prevalence of obesity among adults in Europe rose from 15% to 26% between 1990 and 2022 (World Health Organization, Figure 3).¹ This rise is even more pronounced among children, where the prevalence has doubled, increasing from 5% to 10%. In 2022, around two-thirds of adults and one-fourth of children were overweight.

Figure 3: Evolution of the prevalence of obesity and overweight



Notes: This figure plots the evolution of obesity and overweight prevalence among adults, adolescents, and children from 1990 to 2022. Adults are defined as individuals aged 18 and older, adolescents as those aged 10 to 19, and children as those aged 5 to 9. For adults, overweight is defined as having a body mass index (BMI) over 25, and obesity as having a BMI over 30. In children and adolescents, overweight refers to a BMI of more than one standard deviation above the median, while obesity is defined as a BMI more than two standard deviations above the median. By definition, the "overweight" category includes individuals classified as obese. The estimates come from the Global Health Observatory of the World Health Organization, NCD Risk Factors: BMI.

¹For adults, overweight is defined as having a body mass index (BMI) over 25, and obesity as having a BMI over 30. In children and adolescents, overweight refers to a BMI of more than one standard deviation above the median, while obesity is defined as a BMI of more than two standard deviations above the median. By definition, the "overweight" category includes individuals classified as obese.

It has become a major public health concern, as obesity has serious health implications, increasing the risk of chronic diseases such as cardiovascular disease, type 2 diabetes and cancer. The consequences for children are particularly worrying due to the long-term and persistent negative effects of childhood obesity on health and development. The main causes of obesity and overweight conditions include excessive food intake, lack of physical activity, as well as genetic and environmental factors. In this dissertation, I focus on some policy tools targeting excessive food consumption and more specifically sugar intake.

Drawing on public economics concepts, we see that obesity and overweight create both internalities and externalities. First, overweight and obese individuals bear the costs of their consumption themselves (through reduced life expectancy and increased healthcare expenditures) that they may not correctly internalize. Indeed, they may be imperfectly informed about the health risks or exhibit time-inconsistent behavior, underestimating the future value they will place on avoiding health-related costs [[Allcott et al., 2019b](#)]. Second, sugar consumption imposes costs on society, not just on consumers, as the financial burden on healthcare systems caused by treating conditions linked to sugar consumption is shared by others [[Bhattacharya and Sood, 2011](#)].

To address the obesity epidemic, many European countries are implementing regulations and fiscal policies aimed at both improving consumer information and modifying the environment in which people make their choices.² These measures can be grouped into three main policy approaches [[Lobstein and Neveux, 2021](#)]. The first is fiscal measures (*e.g.*, taxes on food and non-alcoholic beverages), designed to increase the price of unhealthy products and influence consumer decisions toward healthier alternatives. For example, France implemented an excise tax on sugar-sweetened beverages in 2012, which was replaced by a tax based on the added sugar content in 2018 [[Bonnet and Réquillart, 2013](#); [Capacci et al., 2019](#)]. The second approach involves food labeling regulations (*e.g.*, front-of-pack nutritional information) to provide consumers with more information and raise awareness about the products they buy. For example, France launched the Nutri-Score in 2017, a color-coded label that helps consumers to

²While our focus is on regulations and fiscal policies, other interventions, such as awareness campaigns promoting balanced diets and physical activity, are also being implemented to tackle obesity.

quickly assess a product's nutritional quality, from green (A) for healthy to red (E) for less healthy [Dubois et al., 2021]. Lastly, food advertising restrictions have been implemented, such as the EGalim Law in France, which limits advertising for unhealthy foods during children's television programs since 2019.

3.2 The challenges of nutritional taxes

In this dissertation, I focus on the implementation of nutritional taxes. The most prominent example are the taxes targeting the sugar content of non-alcoholic sugar-sweetened beverages (SSB hereafter). In the UK for example, SSBs represent a significant source of sugar intake (22% for children, 33% for adolescents, and 21% for adults [Public Health England, 2018]). To date, more than 50 countries have implemented taxes on sugar-sweetened beverages (Obesity Evidence Hub [2022], see Table 1 for Europe). In France, a volume based tax was first established from 2012 to 2018, and then replaced by a tax with a sliding scale design based on added-sugar content. In contrast, a two tiered sugar-concentration-based tax design was implemented in the United Kingdom based on total sugar content of products.³

Table 1: Sugar taxes on SSBs across Europe

Country	Implemented	Tax design
Belgium	Jan 2016	€0.068/L on soft drinks with added sweeteners €0.41/L and €0.68/100 kg on liquid and powder concentrates
UK	Apr 2018	£0.18/L (5–8 g sugar/100 mL); £0.24/L (>8 g sugar/100 mL) Exempts: dairy drinks, 100% fruit/vegetable juices
Ireland	May 2018	€0.20/L (5–8 g/100 mL); €0.30/L (>8 g/100 mL)
France	Jan 2012 Updated 2018	€0.07/L on drinks with added sugars/artificial sweeteners Sliding scale up to €0.2/L (>11 g/100 mL)
Spain	Jan 2021	21% VAT (increased from 10%) on drinks with added sweeteners/sweetening additives. Exempts: dairy
Catalonia	May 2017	€0.08/L (5–8 g/100 mL); €0.12/L (>8 g/100 mL)
Portugal	Feb 2017	€0.01/L (<2.5 g/100 mL); €0.06/L (2.5–5 g/100 mL); €0.08/L (5–8 g/100 mL); €0.20/L (>8 g/100 mL)
Finland	1940 Updated 2011	€0.22/L on sugar-containing soft drinks €0.12/L on sugar-free soft drinks
Latvia	May 2004 Updated 2016	€0.074/L on drinks with added sugar, sweeteners, or flavoring Exempts: fruit juices (<10% added sugar) and flavored waters without sugar/sweeteners/flavorings
Poland	Jan 2021	PLN 0.5/L + 0.05/g sugar (>5 g/100 mL) on drinks with sweeteners PLN 0.09/L for caffeine/taurine drinks
Hungary	2011	HUF 7/L on soft drinks; HUF 200/L on syrup concentrates

Notes: This table summarizes the design of taxes on SSBs in Europe. Source: UNC, Carolina Population Center.

³France introduced its sugar tax in 2012, which initially applied a rate of 0.07 euros per liter on sugary drinks containing added sugars or sweeteners; and was revised in 2018, to tax drinks based on their added sugar content using a step-wise approach. In the United Kingdom, the Soft Drinks Industry Levy, implemented in 2018, taxes drinks with over 5 grams of sugar per 100 milliliters at 18 pence per liter and those with over 8 grams per 100 milliliters at 24 pence per liter.

The economic rationale for these taxes stems from the Pigouvian concept of externality-correcting taxes, *i.e.*, correcting market inefficiencies by internalizing the negative externalities and thus aligning the private cost with the true social cost. Their goal is to reduce demand by raising prices while generating government revenue. In practice, manufacturers bear the additional cost of production and then decide how much of the tax will be passed on to consumers.⁴ Anticipating the pass-through of the tax is not straightforward because firms can either decide to absorb a part of the tax (the final price increases by less than the tax amount), or fully pass on the tax (the final price increases by the amount of the tax), or to increase their margins (the final price increases by more than the tax amount). Consumers will then adjust their consumption based on the price increase, depending on their sensitivity to price changes (price elasticity of demand). Economic theory predicts that if the price of taxed products rises, consumption will decrease, leading consumers to shift toward relatively cheaper and potentially healthier alternatives not subject to the tax. They can either deter consumption of a whole family of products or encourage substitutions towards healthier products in the same family.

However, the rationale for public intervention, particularly through "sin taxes", is complex [Allcott et al., 2019b; Bhattacharya and Sood, 2011]. On the one hand, these taxes can improve public health and generate government revenue. On the other hand, they are often criticized for being paternalistic and regressive, as these products are disproportionately consumed by lower-income individuals. The effectiveness of such taxes depends on careful design, particularly in targeting the right products and consumers [Allcott et al., 2019a; Dubois et al., 2020; Griffith et al., 2019] and accounting for market power [O'Connell and Smith, 2024]. The assessment of nutritional policies can be based on several criteria such as added sugar consumption, obesity status, tax revenues or welfare impacts.⁵

⁴Manufacturers may also react by changing the product composition. However, the literature mostly focuses on price changes rather than quality changes following the introduction of a tax (see [Allais et al., 2024] for an example on reformulation of product recipes).

⁵In a policy report, commissioned for the Science and Technology in childhood Obesity Policy project, we assessed the potential impacts of new fiscal policies on added sugar in Europe [available here]. We evaluated various hypothetical tax designs and their effects on non-alcoholic beverage and biscuit purchases, as well as sugar intake from these products, particularly among children and adolescents. We analyzed household scanner data to propose ex-ante evaluations and comparisons of the effects of different tax designs in the United Kingdom, France, and Spain. Our findings emphasized the importance of targeting 'unhealthy' food categories with high price sensitivity and appropriately setting tax levels based on the distribution of purchases relative to sugar

In Chapter 3, we develop a new framework for evaluating the optimality of tax design in terms of overall welfare (integrating the social cost of sugar).

4 Methodology of the Thesis

This thesis, based on an empirical approach, draws on a rich set of data (administrative data from firms and individuals, surveys, and household purchase data) and various methods.

4.1 Data sources

Epidemiological cohort. Chapter 1 and 2 exploit the *Constances* epidemiological cohort, a panel of around 450,000 individuals representative of French private sector employees aged 18-69 at inclusion (see [Zins et al. \[2015\]](#)). Individual healthcare utilization are provided by the National Health Insurance Fund with administrative reimbursement records on doctor visits, hospitalizations, dispensed drugs, sick leaves and records of long-term conditions (from 2007 to 2019). Labor market histories are provided by the National Retirement Insurance Fund and include employment status, earnings, occupation and workplace identifier (from 1950 to 2019). The individual data are matched to firms' administrative data thanks to the the workplace identifier. Local contextual indicators are added based on residence at the *commune* or *département* level (local unemployment rate, labor market tightness measure and local healthcare supply).

Firm sources. Chapters 1 and 2 use the matched employer-employee dataset (*Déclarations Annuelles des Données Sociales, DADS - Postes*), which provides detailed job-spell information for each worker, establishment, and year, derived from mandatory fiscal reports (2002-2019). This dataset is used to gather firm characteristics (such as number of employees and occupational structure) and to identify establishment closures. Chapter 2 employs administrative balance sheet data (*FICUS-FARE* and *Bénéfices Réels Normaux, BRN*), which includes financial data and income statements of French firms, collected from mandatory tax reports (2013-2019). This data is used to

content [[Allais et al., 2022](#)].

extract firm characteristics (such as industry, turnover, and value added) and to compute investment in industrial equipment and machines. Chapter 2 also makes use of French customs datasets, which provide comprehensive transaction-level data on the values and quantities of imports by product and country of origin, disaggregated by firm and year (1993-2019). The detailed 8-digit product codes from the Combined Nomenclature (CN-8) are used to identify imports of automation goods and robotics.

Food purchase data. Chapter 3 exploits a representative consumer panel data from Kantar Worldpanel in the United Kingdom (2017). The dataset includes daily home-scan records of food purchases of around 24,500 households. It provides detailed information on households (such as the socio-economic status and the age, weight, height of each member) and product characteristics (such as brand, size, quantity, price, sugar content).

4.2 Empirical methods

Difference-in-differences. Chapters 1 and 2 rely on quasi-experimental designs to estimate causal effects of careers' events on workers' health. In Chapter 1, I leverage plant closures to identify exogenous job losses. In Chapter 2, I detect the adoption of automation technologies at the workplace using i) variations in the stock of industrial equipment and machines, and ii) imports of automation goods and industrial robots. In both chapters, I implement a difference-in-differences (DiD) strategy, comparing changes in outcomes over time between a treated group and a carefully matched control group. This method isolates the impact of the event by measuring the difference in outcomes before and after for both groups and removing any time effects or confounding factors that might have influenced both groups equally. The key assumption underlying this approach is the parallel trends assumption, meaning that in the absence of the event, the treated and control groups would have experienced similar outcome trajectories. I incorporate recent advancements in the DiD literature [[de Chaisemartin and D'Haultfoeuille, 2020a](#); [Goodman-Bacon, 2021](#)] and apply alternative estimators [[Callaway and Sant'Anna, 2021](#); [Cengiz et al., 2019](#); [de Chaisemartin and D'Haultfoeuille, 2020b](#); [Sun and Abraham, 2021](#)] to ensure robustness.

Ex-ante evaluation & structural estimation. The methodology in Chapter 3 builds on the counterfactual evaluation framework from structural industrial organization, as introduced by seminal works of [Berry et al. \[1995, 2004\]](#). This analysis is enriched by incorporating a regulatory framework where total welfare is composed of consumer surplus, firm profits, tax revenue, and the external costs of excessive sugar consumption. These externalities include the healthcare costs associated with conditions like weight gain, type 2 diabetes, and cardiovascular disease [[Allcott et al., 2019b](#)]. By accounting for these costs, the welfare analysis captures the potentially divergent effects on consumers, firms, government revenue, and public health. To estimate the four components of welfare, we employ a three-step structural econometric approach following [Bonnet and Réquillart \[2013\]](#). First, we estimate a discrete choice demand model of food purchases (random coefficient logit model) that captures substitutions within and between varieties of non-alcoholic beverages, identifying household-specific preferences and demand curves. Second, we model the supply side as an oligopoly with differentiated products, where firms compete in a Bertrand-Nash framework, akin to the methods in [Berry et al. \[1995\]](#) and [Nevo \[2001\]](#). This step allows us to estimate price-cost margins and production costs for the tax pass-through. Lastly, we simulate counterfactual scenarios to assess the welfare impacts of regulatory interventions.

5 Summary of the Chapters and Contributions

5.1 Chapter 1 - Job Loss and Health Outcomes: Evidence from France

Motivation. The first chapter explores the immediate and long-term health effects of job loss. In many countries, workers are protected against income loss due to unemployment, yet limited attention is given to the potential negative health impacts. This lack of consideration is surprising, given the significant changes that go along with job loss. The reduction in earnings can lead to stress from economic insecurity and, in some cases, result in the forgoing of healthcare. Additionally, job loss is often associated with lifestyle changes, such as increased isolation and the loss of a structured routine. When comparing the health status of unemployed and employed individuals,

one can observe that the unemployed are more likely to be in poor (mental) health. For example, they are 48% more likely to use antidepressants and anxiolytics. This raises the question of whether these health issues are a cause or a consequence of job loss. On the one hand, deteriorating health might lead to unemployment; on the other hand, the event of job loss itself could contribute to poor health. This chapter investigates whether job loss incurs hidden health costs, with a focus on France, where unemployment benefits are generous and health insurance is universal.

Data & Methods. I use the French epidemiological cohort *Constances*, a unique individual-level dataset that integrates several sources of information. It includes administrative data on labor market history (such as employment, earnings and workplace identifier), healthcare use (such as doctor visits, hospital stays, and dispensed drugs), and survey data on subjective health and risky behaviors. The main empirical challenge in answering this research question is isolating job losses that are not driven by the workers' health. It is crucial to exclude individuals who became unemployed due to pre-existing health conditions. Therefore, the focus is on finding a shock that leads to job termination that is unrelated to workers' health. To achieve this, I use the identification strategy proposed in the seminal paper of [Jacobson et al. \[1993\]](#), which examines mass layoffs and plant closures. I detect these events using administrative matched employer-employee datasets. In the case of plant closures, all workers lose their jobs regardless of their health status, resulting in a group of workers who experienced involuntary displacement, referred to as "treated workers". I compare them to a control group of workers who share similar characteristics but did not face job displacement. Specifically, I compare the evolution of health outcomes in both groups using a difference-in-differences setting. The key assumption is that, in the absence of job loss, the health trajectories of the two groups would have been similar. This allows me to estimate the counterfactual health outcomes for the treated group, showing what their health would have been had they not lost their jobs.

Results & Contributions. The results reveal significant long-term economic and health impacts of job displacement. Displaced workers experience a sharp decline in annual

earnings, losing 5,000 € in the first year, with earnings still 3,000 € lower five years later. Mental health is notably affected, with antidepressant and anxiolytic use rising by 19% and sleeping pill use increasing by 32%, both in the short and long term. Displaced workers are also more likely to consult psychiatrists and experience hospitalizations for mental health issues shortly after job loss. Physical health deteriorates as well, with increased doctor visits, hospital admissions, and a more than doubling of the probability of receiving disability benefits. These effects are particularly severe among vulnerable groups, including older workers, low-skilled workers, and those living in high-unemployment regions. I show that the mechanism of new job opportunities plays a crucial role, as the increase in mental healthcare use is concentrated among high-risk workers, *i.e.*, those with a higher risk of unemployment. The mental health decline is mostly driven by the immediate shock of job loss, rather than by the duration of unemployment, with antidepressant use peaking right after displacement.

5.2 Chapter 2 - Automation at the Workplace: Implications for Workers' Health

This chapter is co-authored with Pauline Lesterquy (Banque de France, CREST).

Motivation. In recent centuries, historians have documented successive waves of industrial revolutions, with technological progress profoundly reshaping employment and job content. This chapter focuses on the most recent two waves, which have been particularly characterized by the advent of automation and robotics. Specifically, we aim at understanding the impact of adopting these innovations in the workplace on workers' health. A large body of literature has documented the direct effects of these technological advances on the nature of work and the tasks performed by workers. The negative perspective emphasizes the substitution of automated machines for human labor, leading to job displacement [[Acemoglu and Restrepo, 2020](#); [Autor et al., 2003](#); [Bessen et al., 2023](#); [Goos, 2018](#)]. In contrast, the positive view highlights the productivity gains brought about by technological change [[Aghion et al., 2024](#); [Bonfiglioli et al., 2024](#); [Graetz and Michaels, 2018](#)]. These transformations in production processes may also influence workers' well-being. However, the overall impact on workers' health remains under-explored. The expected effects are ambiguous. In terms of mental health,

the risk of displacement can induce stress, yet the introduction of more fulfilling tasks could potentially increase job satisfaction. Regarding physical health, improved working conditions, such as less physically demanding tasks, might benefit workers, though the risk of workplace injuries from working alongside machines could rise. Additionally, changes in work intensity may affect both mental and physical health. The chapter empirically quantifies the average effects and also look at sub-groups who could be affected differently.

Data & Methods. We focus on the manufacturing sector in France between 2007 and 2019. For the first set of analyses at the firm-level, we use different firm-level administrative panel datasets (customs, balance-sheets and matched employer-employee dataset) to capture firm adoption of industrial equipment and automation technologies. Since direct observation of automation adoption is challenging, we rely on two key proxies. The first proxy uses balance-sheet data on industrial equipment and machines to estimate firms' investments in physical capital. The second proxy focuses on firm-level imports of automation goods (or alternatively industrial robots), derived from detailed customs data. We detect automation events by considering large investments in manufacturing capital goods using these two proxies. We aim to capture shifts in production processes that result from automation adoption. For the second set of analyses at the worker-level, we exploit administrative records of individual healthcare from the French epidemiological cohort *Constances*. This individual dataset is matched with the firm datasets with the workplace identifier. This allows us to identify workers exposed to firm-level technological change. The rich dataset on healthcare use allows us to study the effects of automation on both mental and physical health and to investigate the potential mechanisms. We also rely on difference-in-differences methodology where we compare a group of treated firms (or workers) to a group of control firms (or workers), with the identifying assumption that their outcome would have evolved in the same way had there been no adoption of automation technology.

Results & Contributions. First, we characterize what happens when a firm initiates a technological shift. We observe that firms increase the number of employees with com-

positional effects because the share of blue collar workers increases. This is responsible for a decrease in the average hourly wage. We also observe an increase in labor productivity and a decrease in the labor share. The results support the positive view of automation, meaning that in our setting the productivity channels outweighs the displacement channel at the firm level on average. Second, we consider the consequences on workers outcomes. Incumbent workers do not experience significant changes in their labor market outcomes (earnings and employment status) on average but older workers face a high displacement risk. We document a deterioration in mental health, with an average 27% increase in the use of antidepressants and anxiolytics. Physical health remains unaffected overall, except for a decrease in anti-inflammatory drug use among low-wage workers. Furthermore, both older workers and blue-collar workers show a marked rise in disability benefits claims, while older workers, in particular, take an additional six days of sick leave on average.

5.3 Chapter 3 - Are Sugar Taxes Well Designed? Empirical evidence from the UK soft drink market

This chapter is co-authored with Olivier Allais (INRAE, PSAE), Céline Bonnet (INRAE, TSE) and Maxime Tranchard (INRAE, PSAE).

Motivation. Policymakers are increasingly turning to nutritional taxes as a tool to steer consumer behavior toward healthier diets. So far, over 50 countries have introduced taxes on sugar-sweetened beverages (SSBs). However, the efficiency of these taxes largely depends on their design. Both theoretical and empirical research suggest that taxes should ideally align with the degree of harm caused by consumption. Yet, in practice, none of the nutritional taxes implemented so far follow a strictly linear structure (tax rate proportional to sugar content). Instead, most feature tiered designs with thresholds that differ in their proximity. Indeed, real-world implementations might differ from economic theory due to feasibility constraints. This raises the question: can such tiered taxes still be optimal, and if so, how should their key components be structured?

Data & Methods. We use home-scan consumer panel data on the non-alcoholic bev-

erage market in the United Kingdom for 2017. This representative panel provides detailed information on household characteristics and product attributes, including brand, size, quantity, price, and sugar content. We develop a regulatory framework in which the objective function is the total welfare, encompassing consumer surplus, firm profit, fiscal revenue, and the social cost of sugar (represented by healthcare externalities due to excessive sugar consumption). The regulator wants to select the tax design that maximizes total welfare based on three parameters: the number of thresholds, the location of the thresholds (in g/100mL), and the tax rates (in euro cents/L). This flexible setting enables the exploration of a wide range of tax structures, from proportional sugar-content taxes to tiered designs with varying numbers of thresholds. We rely on the estimate of the social cost of sugar by [Rischbieth et al. \[2020\]](#), at approximately 27 € per kilogram of sugar consumption. We assess the effects of nutritional fiscal policies using ex-ante structural econometric methods that capture both consumer and firm reactions to tax changes. Following the three-step methodology outlined by [Bonnet and Réquillart \[2013\]](#), we first estimate demand using a random coefficient logit model to understand consumer preferences for price and quality. Next, we model the supply side as an oligopolistic competition, allowing us to analyze pricing strategies, marginal costs, and the pass-through of taxes by firms. Finally, we conduct counterfactual simulations of various taxation scenarios to study their effects on total welfare.

Results & Contributions. Our findings indicate that tiered tax structures, which consider firms' strategic responses, can deliver significantly greater welfare gains than the UK's current Soft Drinks Industry Levy (SDIL). In particular, the optimal design imposes higher taxes on high-sugar products, encouraging firms to lower the prices of lower-sugar alternatives. This leads not only to an improvement in public health outcomes but also to an increase in consumer surplus while maintaining firm profitability. These results suggest that policymakers should carefully factor in firms' strategic behavior when designing sugar taxation policies to maximize welfare benefits.



Introduction générale

La santé des individus dépend d'un ensemble varié de facteurs interconnectés, allant d'éléments biologiques à des conditions sociales et environnementales plus larges. Certains de ces déterminants restent inexpliqués ou échappent au contrôle des individus, comme les prédispositions génétiques et les dynamiques sociales complexes. Cependant, d'autres causes ont été bien identifiées, notamment les conditions de travail, les niveaux de revenu, l'accès aux soins de santé, ainsi que les habitudes de vie telles que l'alimentation et l'activité physique, et leurs effets ont été largement étudiés. Les économistes jouent un rôle crucial dans ce contexte : en analysant les causes prévisibles et mesurables des problèmes de santé, ils peuvent proposer des politiques publiques visant à améliorer le bien-être de la population, en particulier en s'attaquant aux inégalités et en réduisant les risques associés à certains environnements.

Cette thèse se situe à l'intersection de l'économie de la santé avec deux domaines : le marché du travail et l'organisation industrielle. L'objectif est d'approfondir notre compréhension de deux facteurs clés qui influencent à la fois la santé individuelle et publique. Tout d'abord, j'explore comment certains événements de carrière, tels que la perte d'emploi (Chapitre 1) et l'automatisation au travail (Chapitre 2), impactent la santé physique et mentale des travailleurs. Ensuite, j'examine la conception des taxes sur le sucre et ses implications pour le bien-être global (Chapitre 3).

Cette introduction est structurée en cinq sections. La première section aborde les défis associés à la mesure de la santé individuelle. La deuxième section explore l'impact du marché du travail sur la santé des travailleurs. La troisième section décrit l'épidémie d'obésité et se concentre sur une réponse politique : les taxes nutritionnelles. La qua-

trième section fournit un aperçu des sources de données et des méthodes empiriques utilisées. Enfin, la cinquième section résume chaque chapitre en mettant en avant leurs contributions à la littérature existante.

1 Mesurer la Santé Individuelle

L'Organisation Mondiale de la Santé (OMS) définit la santé comme "un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité". Cette définition souligne que la santé est une notion complexe, façonnée par une combinaison d'expériences et de perceptions. Il existe plusieurs façons de mesurer les indicateurs de santé dans les données, chaque méthode offrant ses propres avantages et limites.

1.1 Santé objective et subjective

La santé objective fait référence à des mesures quantifiables telles que des indicateurs cliniques (par exemple tension artérielle, taux de cholestérol, indice de masse corporelle), des diagnostics de maladies (par exemple diabète, affections cardiaques) ou des indicateurs de santé comme les taux de mortalité et d'hospitalisation. Ces mesures offrent une perspective claire et standardisée sur la santé, généralement obtenues à partir de données administratives (par exemple visites médicales, hospitalisations, médicaments prescrits), de diagnostics médicaux, de biomarqueurs ou d'évaluations physiques.

En revanche, la santé subjective repose sur les perceptions et les auto-évaluations des individus concernant leur bien-être. Ces données sont généralement recueillies par des enquêtes ou des questionnaires où les individus évaluent leur propre santé, état mental ou fonctionnement quotidien. Les outils courants incluent des questions sur la santé auto-évaluée ("Comment évalueriez-vous votre santé globale ?") et des échelles de bien-être mental (par exemple niveaux de stress, d'anxiété ou de satisfaction de vie). Bien que les mesures subjectives puissent manquer de la précision des données cliniques, elles offrent des informations précieuses sur la perception que les individus ont de leur santé, englobant des impacts plus larges sur la qualité de vie. De plus, les

données de santé subjective permettent l'étude d'une sélection différente d'individus, car elles ne se limitent pas à ceux qui consultent des professionnels de santé.

1.2 Santé mentale et physique

La santé physique est généralement évaluée à l'aide d'indicateurs objectifs tels que la tension artérielle, l'indice de masse corporelle (IMC), le taux de cholestérol ou la présence de maladies chroniques comme le diabète ou les affections cardiovasculaires. Ces indicateurs sont couramment extraits des dossiers médicaux, des examens de santé ou des données administratives.

La santé mentale est plus difficile à mesurer objectivement et est souvent évaluée par des enquêtes auto-déclarées ou des évaluations psychologiques (par exemple des échelles standardisées comme le GHQ-12 ou le PHQ-9) qui abordent les symptômes de dépression, d'anxiété ou de bien-être émotionnel général. Cependant, le recours aux soins, comme les prescriptions d'antidépresseurs, peut fournir des informations supplémentaires sur les problèmes de santé mentale actuels ou passés.

Bien qu'elles soient mesurées séparément, la santé physique et mentale sont étroitement liées, car les problèmes de santé mentale peuvent influencer les indicateurs de santé physique, et les affections physiques chroniques peuvent avoir un impact significatif sur le bien-être mental.

1.3 État de santé et consommation de soins

Les consommations de soins font référence au recours aux services médicaux, souvent enregistrée dans les données administratives. Bien que la consommation de soins de santé puisse refléter les besoins de santé et soit fréquemment utilisée comme indicateur de la santé, elle ne capture pas toujours précisément l'état de santé d'un individu. Par exemple, une forte consommation d'antidépresseurs pourrait indiquer une amélioration de la santé mentale grâce à un traitement efficace, mais elle pourrait aussi signaler des problèmes de santé mentale persistants nécessitant une médication continue. Inversement, une faible consommation d'antidépresseurs ne signifie pas nécessairement une bonne santé mentale, car des obstacles comme l'accès aux soins, la stigmatisation

ou des contraintes financières peuvent empêcher les individus de rechercher un traitement nécessaire.

Dans cette thèse, je m'appuie principalement sur des mesures de santé objectives fournies par les données de remboursement, qui offrent des enregistrements détaillés des visites chez les médecins, des séjours hospitaliers, des médicaments délivrés et des arrêts maladie pour la santé physique et mentale.

1.4 Quelques chiffres pour la France

L'évolution des maladies chroniques a régulièrement augmenté au cours des dernières décennies en France. En 2021, environ 30% de la population adulte française vivait avec au moins une maladie chronique, les maladies cardiovasculaires, le diabète et les troubles respiratoires étant les plus fréquents [[Santé Publique France, 2021](#)]. Environ 20% de la population française éprouve des troubles de santé mentale à un moment donné de leur vie, l'anxiété et la dépression étant les plus courants [[INSERM, 2020](#)]. Les dépenses de santé en France ont constamment augmenté au fil des décennies, atteignant près de 12% du PIB en 2022. Cela fait de la France l'un des pays aux dépenses de santé les plus élevées en Europe. La hausse des coûts est en grande partie attribuée au vieillissement de la population, aux innovations médicales coûteuses et à une demande croissante de soins pour maladies chroniques. La majorité des dépenses est prise en charge par le système national d'assurance maladie, avec une part croissante consacrée aux soins ambulatoires, aux médicaments et aux dépenses hospitalières [[European Observatory on Health Systems and Policies, 2023](#)].

2 L'interaction entre le Travail et la Santé

2.1 L'effet de l'emploi sur la santé

L'emploi et la santé sont intrinsèquement liés tout au long de la vie. Les individus confrontés à des problèmes de santé ont souvent des difficultés à obtenir un emploi stable. Inversement, la nature de l'emploi et les conditions de travail peuvent affecter la santé de manière significative et parfois ambiguë.

L'emploi peut avoir des effets positifs considérables sur la santé. Au-delà d'une source de revenu, il procure aux individus un sentiment d'utilité, une structure et des interactions sociales, contribuant ainsi à un meilleur bien-être mental et émotionnel. En outre, les activités liées au travail encouragent souvent l'activité physique et stimulent les capacités cognitives. De plus, dans certains contextes, les emplois offrent des avantages en matière de santé, comme une couverture d'assurance incluant des dépistages préventifs et des traitements, ce qui peut améliorer la santé physique [Burgard and Lin, 2013; Marmot and Stansfeld, 2003].

Cependant, certaines conditions de travail peuvent avoir un impact négatif sur la santé. Des niveaux de stress élevés, de longues heures de travail et l'insécurité de l'emploi sont associés à des indicateurs de santé mentale détériorés, notamment l'anxiété, la dépression et le burnout [Berniell and Bietenbeck, 2020; Blasco et al., 2022; Caroli and Godard, 2016]. De plus, des environnements de travail physiquement exigeants ou dangereux peuvent entraîner des blessures ou des problèmes de santé chroniques. Les modes de vie sédentaires, courants dans les emplois de bureau, et la difficulté à maintenir un équilibre entre vie professionnelle et vie personnelle peuvent également aggraver les problèmes de santé, contribuant à des affections comme les maladies cardiovasculaires, les troubles musculo-squelettiques et l'obésité.

Je vise à éclairer l'impact de deux chocs de carrière sur la santé mentale et physique des travailleurs. Dans le Chapitre 1, je me concentre sur la perte d'emploi involontaire, qui représente principalement un choc sur la marge extensive. Dans le Chapitre 2, j'étudie les changements induits par l'automatisation dans les conditions de travail, ainsi que la destruction ou la réallocation de l'emploi.

2.2 Le choc de carrière de la perte d'emploi

Un choc de carrière commun et de relativement court-terme est la perte d'emploi, qui peut avoir des effets immédiats et durables sur les individus. Ce choc agit sur la marge extensive de l'emploi. En France, environ 2,9 millions de personnes étaient au chômage en 2023, représentant environ 7,2% de la population active. Bien que certains travailleurs se rétablissent rapidement, les conséquences de la perte d'emploi peuvent

persister, affectant à la fois le revenu et le bien-être général. La perte d'emploi a des conséquences économiques significatives, notamment par la réduction des revenus, pouvant entraîner une instabilité financière à long terme. Les travailleurs licenciés subissent des pertes de revenus persistantes, avec des revenus souvent inférieurs même plusieurs années après la perte d'emploi [Jacobson et al., 1993]. Beaucoup de travailleurs peinent à retrouver les niveaux de salaire d'avant leur licenciement, ce qui accroît leur vulnérabilité économique.

Au-delà du revenu, la perte d'emploi entraîne des changements de mode de vie majeurs, notamment une augmentation du stress, de l'incertitude et des modifications dans la dynamique familiale, pouvant entraîner une diminution de l'engagement social et des changements dans les habitudes de consommation. Ces bouleversements peuvent contribuer à un déclin de la santé mentale et physique, car il a été démontré que la pression financière et le stress augmentent le risque de dépression, d'anxiété et même de maladies chroniques comme les maladies cardiaques [Sullivan and von Wachter, 2009].

2.3 La nature changeante du travail : avancées technologiques et leurs impacts

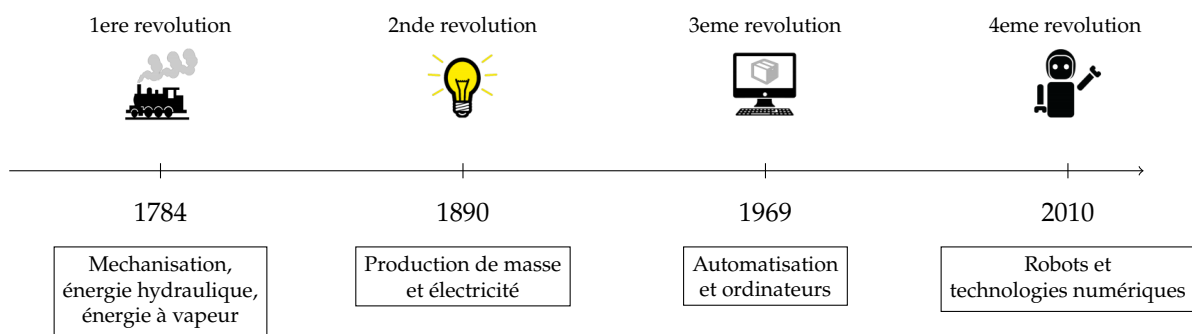
Le marché du travail a connu d'importants changements structurels ces dernières années, principalement en raison de l'innovation technologique. Ces progrès ont non seulement redéfini la nature du travail, mais ont également modifié la composition et la complexité des tâches.

2.3.1 Un bref historique du changement technologique

La chronologie des révolutions industrielles illustrée dans la Figure 4 montre quatre grandes phases de transformation technologique et industrielle qui ont façonné les économies modernes. La première révolution industrielle ("révolution de la machine à vapeur"), débutant en 1784, est caractérisée par l'introduction de la mécanisation, l'utilisation de l'eau et de la vapeur, et l'émergence des premières usines. Elle a ouvert la voie à la production de masse et de l'utilisation généralisée de l'électricité lors de la seconde révolution industrielle ("révolution de l'électricité") à la fin du 19^{ème} siècle. En 1969, la troisième révolution industrielle ("révolution de l'automatisation") a

émergé avec l'essor de l'automatisation et l'apparition des ordinateurs, changeant fondamentalement la nature du travail et de la production. Enfin, la quatrième révolution industrielle ("révolution de la digitalisation"), débutant en 2010, est marquée par les avancées en robotique et en technologies numériques, qui continuent de repousser les limites de l'automatisation et de la connectivité dans l'économie mondiale. Dans cette thèse, je me concentre sur les deux révolutions industrielles les plus récentes et sur l'adoption de l'automatisation, de la robotique et de la digitalisation sur les lieux de travail.

Figure 4: Frise chronologique des révolutions industrielles

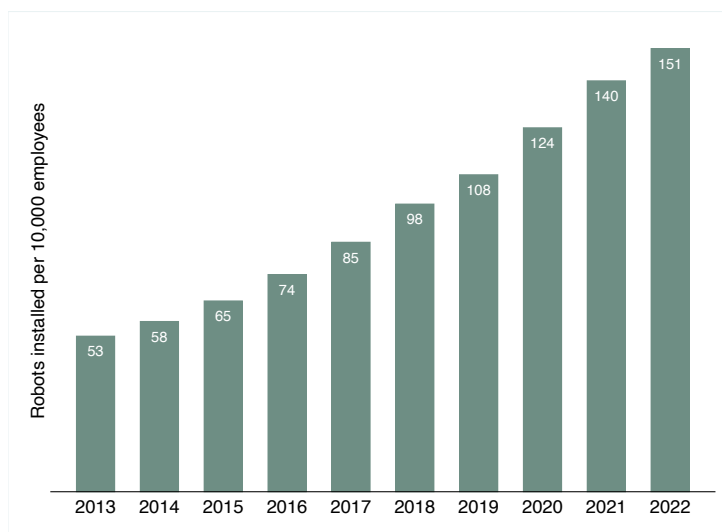


2.3.2 Comment la main-d'œuvre sera-t-elle affectée ?

Ces dernières années, la pénétration des robots industriels dans les entreprises a considérablement augmenté dans le monde. La densité de robots, qui s'établissait à 53 robots pour 10 000 employés en 2013, est montée à 151 en 2022 (Figure 5). Cette forte augmentation souligne la dépendance croissante à l'automatisation dans divers secteurs et suggère que les travailleurs sont de plus en plus exposés à sa présence sur le lieu de travail.

L'adoption de ces innovations dans le milieu professionnel est susceptible d'affecter une part importante de la main-d'œuvre. Presque la moitié de tous les emplois sont menacés d'automatisation aux États-Unis [Brynjolfsson and McAfee, 2014; Frey and Osborne, 2017] et jusqu'à 60% des emplois dans les pays de l'OCDE [World Bank, 2016]. Une enquête récente en France indique que plus d'un tiers des salariés travaillent à un rythme surveillé par des ordinateurs en France [DARES, 2021].

Figure 5: Évolution de la densité de robots



Notes : Cette figure montre l'évolution de la densité de robots de 2013 à 2022 dans le monde. La densité de robots est calculée comme le nombre de robots industriels installés pour 10 000 employés. Source : World Robotics 2023, Fédération Internationale de la Robotique.

L'oeuvre influente de [Autor et al. \[2003\]](#) fournit des aperçus précieux sur la manière dont ces innovations peuvent redéfinir la nature du travail et modifier la demande pour différents types de travail. Ils classent les tâches selon deux dimensions clés : cognitives vs manuelles, et routinières vs non routinières. Les tâches cognitives nécessitent la résolution de problèmes et la prise de décisions complexes, tandis que les tâches manuelles impliquent un travail physique. Les tâches routinières sont répétitives et suivent un ensemble d'étapes prédéfinies, tandis que les tâches non routinières sont plus complexes et nécessitent de la flexibilité, de la résolution de problèmes et de l'adaptabilité. Par nature, les tâches routinières sont plus susceptibles d'être automatisées. Par exemple, les ordinateurs sont plus susceptibles de remplacer les travailleurs effectuant des tâches routinières et codifiables, tout en servant de complément et en améliorant les capacités des travailleurs pour les tâches cognitives non routinières (comme la résolution de problèmes et la créativité) et les tâches manuelles (comme celles nécessitant de l'adaptabilité en fonction de la situation) [[Autor et al., 2003](#)]. Le modèle des tâches développé dans [Acemoglu and Autor \[2011\]](#) rationalise ce changement technologique orienté vers les tâches routinières et la manière dont les travailleurs sont affectés différemment en fonction de leurs compétences initiales.

Cette décomposition aide à comprendre le phénomène de la polarisation des em-

plis [Autor and Dorn, 2013; Goldin and Katz, 1998; Goos, 2018; Goos and Manning, 2007]. Ces dernières décennies, une tendance notable a été la polarisation du marché du travail, où les augmentations salariales ont disproportionnellement favorisé les travailleurs en haut et au bas de la distribution des revenus et des compétences, laissant ceux du milieu derrière. Les avancées technologiques peuvent être une des raisons de ce phénomène. Par exemple, le capital numérique a un avantage comparatif dans l'exécution des tâches routinières qui peuvent être facilement codifiées dans des logiciels, évincant ainsi les travailleurs de compétences intermédiaires. En conséquence, certains travailleurs de compétences intermédiaires réallouent leur travail vers des tâches plus simples, tandis que d'autres passent à des rôles plus complexes.

2.3.3 Les travailleurs doivent-ils s'inquiéter de l'automatisation et des technologies numériques ?

La propagation rapide de l'automatisation et des nouvelles technologies a suscité de vives inquiétudes parmi les travailleurs, alimentant la crainte que le travail humain ne devienne obsolète. Après la "révolution de l'électricité", Keynes écrivait en 1930 *"Nous sommes affligés par une nouvelle maladie dont certains lecteurs n'ont peut-être pas encore entendu parler, mais dont ils entendront beaucoup parler dans les années à venir – à savoir le chômage technologique"*. En 1961, le magazine TIME publiait un article intitulé "The Automation Jobless", avertissant du potentiel de chômage massif dû aux avancées technologiques [Autor, 2015]. Ces craintes de licenciements généralisés surgissent lorsque l'automatisation opère sur la marge extensive, créant des effets de substitution à mesure que les technologies prennent en charge des tâches auparavant effectuées par des travailleurs à moindre coût. Cela conduit à la destruction d'emplois et à une baisse de la demande de travail, exerçant finalement une pression à la baisse sur les salaires [Acemoglu and Restrepo, 2020; Bessen et al., 2023; Graetz and Michaels, 2018]. Par exemple, considérons le secteur de la fabrication automobile, où les systèmes robotiques ont largement remplacé le travail manuel pour des tâches telles que la soudure, la peinture et l'assemblage. Autrefois effectuées par des travailleurs humains, ces tâches sont désormais automatisées, réduisant ainsi le besoin d'une main-

d'œuvre nombreuse et faisant baisser les salaires pour les rôles restants, plus intensifs en travail humain.

Cependant, l'histoire montre que le travail humain n'est pas devenu obsolète [Autor, 2015]. Malgré les avancées significatives dans l'automatisation des tâches, l'emploi global n'a pas diminué. Plusieurs mécanismes contrebalancent l'effet de la destruction des emplois. Une explication clé est que l'automatisation peut être complémentaire au travail humain. L'effet de productivité joue ici un rôle crucial : à mesure que l'automatisation réduit les coûts de production, elle augmente la demande pour des tâches non automatisées [Acemoglu et al., 2020; Aghion et al., 2024; Bonfiglioli et al., 2024]. De plus, l'automatisation opère souvent sur la marge intensive, améliorant les tâches existantes plutôt que de simplement remplacer le travail humain. Enfin, le mécanisme prévalent est la création de nouvelles tâches, où le travail humain conserve un avantage comparatif sur les machines, favorisant ainsi de nouvelles opportunités d'emploi [Acemoglu and Restrepo, 2019]. Dans l'industrie automobile, par exemple, bien que les robots aient pris en charge les tâches répétitives telles que la soudure ou la peinture, les travailleurs humains sont toujours nécessaires pour la prise de décisions complexes et le contrôle de la qualité, comme dans la conception, l'ingénierie et les inspections, où leur jugement et leurs compétences sont essentiels.

Ces exemples montrent que même lorsque certaines tâches sont automatisées, les travailleurs peuvent encore être nécessaires pour effectuer des tâches complémentaires. Au lieu de devenir obsolètes, les travailleurs peuvent devoir s'adapter en appliquant différentes compétences aux tâches en évolution dans leurs rôles. À mesure que les emplois se transforment, une réallocation du travail se produira : certains travailleurs connaîtront des changements dans leurs postes actuels, tandis que d'autres découvriront que d'autres rôles sont désormais mieux adaptés à leurs compétences. La transition peut être lente et douloureuse, en particulier pour les travailleurs dont les compétences ne correspondent pas aux exigences du nouveau paysage technologique [Acemoglu and Restrepo, 2020; Bessen et al., 2023; Bonfiglioli et al., 2024; Humlum, 2021; Yaky-movych, 2022]. Ce décalage entre la technologie et les compétences constitue un défi majeur, car les travailleurs évincés par l'automatisation ne possèdent pas immédiate-

ment les compétences requises pour les emplois nouvellement créés.

2.3.4 Faut-il s'attendre à des effets de l'automatisation sur la santé des travailleurs ?

Les changements dans les tâches et les conditions de travail peuvent avoir un impact sur la santé physique et mentale des travailleurs, avec des effets qui pourraient varier en fonction des caractéristiques individuelles des travailleurs. Ces transformations devraient influencer la prévalence des tâches physiques, la fréquence des accidents du travail et l'intensité générale du travail. Le sens du travail pourrait être négativement affecté par une surveillance automatisée accrue, ou bien il pourrait s'améliorer si l'automatisation permet aux travailleurs d'effectuer des tâches plus épanouissantes. De plus, des préoccupations liées au stress pourraient surgir en raison de l'intensification des objectifs de production ou de l'anxiété concernant la sécurité de l'emploi, alors que les travailleurs perçoivent une concurrence avec les robots et les systèmes automatisés.

Dans le Chapitre 2, je quantifie les effets de l'automatisation et de la robotique sur le lieu de travail sur la santé des travailleurs et j'examine les mécanismes sous-jacents.

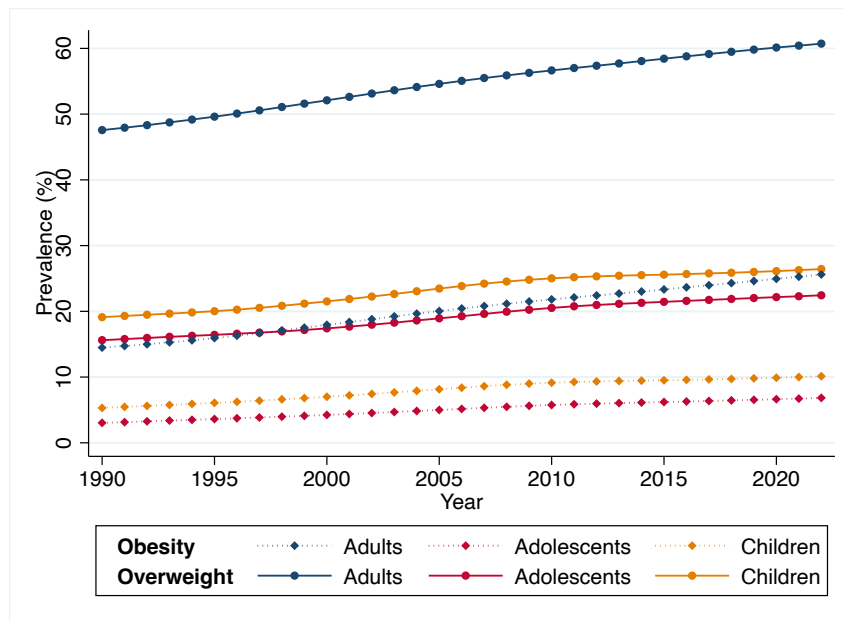
3 Lutter contre l'Épidémie d'Obésité : le cas des Taxes sur le Sucre

3.1 Évolutions de l'obésité et outils politiques

On observe une augmentation alarmante des taux d'obésité et de surpoids dans de nombreux pays ces dernières années et ce, dans toutes les tranches d'âge. La prévalence de l'obésité chez les adultes en Europe est passée de 15% à 26% entre 1990 et 2022 (Organisation mondiale de la santé, Figure 6).⁶ Cette augmentation est encore plus marquée chez les enfants, où la prévalence a doublé, passant de 5% à 10%. En 2022, environ les deux tiers des adultes et un quart des enfants étaient en surpoids.

⁶Pour les adultes, le surpoids est défini comme ayant un indice de masse corporelle (IMC) supérieur à 25, et l'obésité comme un IMC supérieur à 30. Pour les enfants et les adolescents, le surpoids est défini comme un IMC de plus d'une déviation standard au-dessus de la médiane, tandis que l'obésité est définie comme un IMC de plus de deux déviations standards au-dessus de la médiane. Par définition, la catégorie "surpoids" inclut les individus classés comme obèses.

Figure 6: Evolution de la prévalence de l'obésité et du surpoids



Notes: Cette figure représente l'évolution de la prévalence du surpoids et de l'obésité chez les adultes, les adolescents et les enfants entre 1990 et 2022. Les adultes sont définis comme des individus âgés de 18 ans et plus, les adolescents comme ceux âgés de 10 à 19 ans, et les enfants comme ceux âgés de 5 à 9 ans. Pour les adultes, le surpoids est défini par un indice de masse corporelle (IMC) supérieur à 25, et l'obésité par un IMC supérieur à 30. Pour les enfants et les adolescents, le surpoids est défini comme un IMC supérieur d'une déviation standard à la médiane, tandis que l'obésité est définie comme un IMC supérieur de deux déviations standards à la médiane. Par définition, la catégorie "surpoids" inclut les individus classés comme obèses. Les estimations proviennent de l'Observatoire mondial de la santé de l'Organisation mondiale de la santé, Facteurs de risque des MNT : IMC.

Cela est devenu un problème majeur de santé publique, car l'obésité a des conséquences graves sur la santé, augmentant le risque de maladies chroniques telles que les maladies cardiovasculaires, le diabète de type 2 et le cancer. Les conséquences pour les enfants sont particulièrement préoccupantes en raison des effets négatifs à long terme et persistants de l'obésité infantile sur la santé et le développement. Les principales causes de l'obésité et du surpoids incluent une consommation excessive de nourriture, un manque d'activité physique, ainsi que des facteurs génétiques et environnementaux. Dans cette dissertation, je me concentre sur certains outils de politique publique visant à réduire la consommation excessive de nourriture et plus spécifiquement l'apport en sucre.

En s'appuyant sur les concepts de l'économie publique, on peut observer que l'obésité et le surpoids génèrent à la fois des internalités et des externalités. Tout d'abord, les individus en surpoids ou obèses supportent eux-mêmes les coûts de leur consommation (par la réduction de l'espérance de vie et l'augmentation des dépenses de santé), qu'ils peuvent ne pas bien internaliser. En effet, ils peuvent être mal informés des risques

sanitaires ou adopter un comportement temporellement incohérent, sous-estimant la valeur future qu'ils accorderont à éviter les coûts liés à la santé [Allcott et al., 2019b]. Ensuite, la consommation de sucre impose des coûts à la société, et pas seulement aux consommateurs, car la charge financière des systèmes de santé due au traitement des maladies liées à la consommation de sucre est partagée par d'autres [Bhattacharya and Sood, 2011].

Pour lutter contre l'épidémie d'obésité, de nombreux pays européens mettent en œuvre des régulations et des politiques fiscales visant à la fois à améliorer l'information des consommateurs et à modifier l'environnement dans lequel les individus font leurs choix.⁷ Ces interventions publiques peuvent être regroupées en trois approches principales [Lobstein and Neveux, 2021]. La première consiste en des mesures fiscales (par exemple des taxes sur les aliments et boissons non alcoolisées), visant à augmenter le prix des produits mauvais pour la santé et influencer les choix des consommateurs vers des alternatives plus saines. Par exemple, la France a mis en place une taxe sur les boissons sucrées en 2012, qui a été remplacée par une taxe basée sur la teneur en sucre ajouté en 2018 [Bonnet and Réquillart, 2013; Capacci et al., 2019]. La deuxième approche concerne les régulations sur l'étiquetage des aliments (par exemple les informations nutritionnelles sur le devant des emballages) pour fournir aux consommateurs plus d'informations et sensibiliser sur les produits qu'ils achètent. Par exemple, la France a lancé le Nutri-Score en 2017, une étiquette colorée qui aide les consommateurs à évaluer rapidement la qualité nutritionnelle d'un produit, du vert (A) pour les produits sains au rouge (E) pour les moins sains [Dubois et al., 2021]. Enfin, des restrictions sur la publicité alimentaire ont été mises en place, comme la loi EGAlim en France, qui limite la publicité pour les aliments nocifs pendant les émissions télévisées pour enfants depuis 2019.

3.2 Les défis des taxes nutritionnelles

Dans cette dissertation, je me concentre sur la mise en œuvre des taxes nutritionnelles. L'exemple le plus marquant concerne les taxes ciblant la teneur en sucre des bois-

⁷Bien que notre focus soit sur les régulations et les politiques fiscales, d'autres interventions, telles que des campagnes de sensibilisation sur les régimes alimentaires équilibrés et l'activité physique, sont également mises en œuvre pour lutter contre l'obésité.

sons non alcoolisées sucrées (BRSA, pour boissons rafraîchissantes sans alcool). Au Royaume-Uni, par exemple, les BRSA représentent une source importante d'apport en sucre (22% pour les enfants, 33% pour les adolescents et 21% pour les adultes [[Public Health England, 2018](#)]). À ce jour, plus de 50 pays ont mis en place des taxes sur les boissons sucrées ([Obesity Evidence Hub \[2022\]](#), voir la Table 2 pour l'Europe). En France, une taxe basée sur le volume a été instaurée de 2012 à 2018, puis remplacée par une taxe à paliers sur la teneur en sucre ajouté. En revanche, le Royaume-Uni a mis en place une taxe à deux niveaux, fondée sur la concentration en sucre, calculée en fonction du sucre total des produits.⁸

Table 2: Taxes sur les boissons sucrées en Europe

Pays	Mise en œuvre	Conception de la taxe
Belgique	Janv. 2016	0,068 €/L sur les boissons avec édulcorants ajoutés 0,41 €/L et 0,68 €/100 kg sur les concentrés liquides et en poudre
Royaume-Uni	Avril 2018	0,18 £/L (5–8 g de sucre/100 mL) ; 0,24 £/L (>8 g de sucre/100 mL) Exemptions : boissons lactées, jus de fruits/légumes 100%
Irlande	Mai 2018	0,20 €/L (5–8 g/100 mL) ; 0,30 €/L (>8 g/100 mL)
France	Janv. 2012 Mis à jour en 2018	0,07 €/L sur boissons avec sucres ajoutés ou édulcorants artificiels Barème progressif jusqu'à 0,2 €/L (>11 g/100 mL)
Espagne	Janv. 2021	TVA à 21% (augmentée depuis 10%) sur boissons avec édulcorants ou additifs sucrants. Exemptions : produits laitiers
Catalogne	Mai 2017	0,08 €/L (5–8 g/100 mL) ; 0,12 €/L (>8 g/100 mL)
Portugal	Févr. 2017	0,01 €/L (<2,5 g/100 mL) ; 0,06 €/L (2,5–5 g/100 mL) ; 0,08 €/L (5–8 g/100 mL) ; 0,20 €/L (>8 g/100 mL)
Finlande	1940 Mis à jour en 2011	0,22 €/L sur les boissons contenant du sucre 0,12 €/L sur les boissons sans sucre
Lettonie	Mai 2004 Mis à jour en 2016	0,074 €/L sur boissons avec sucres, édulcorants ou arômes ajoutés Exemptions : jus de fruits (<10% de sucre ajouté) et eaux aromatisées sans sucres/édulcorants/arômes
Pologne	Janv. 2021	0,5 PLN/L + 0,05 PLN/g sucre (>5 g/100 mL) sur boissons avec édulcorants 0,09 PLN/L pour les boissons avec caféine ou taurine
Hongrie	2011	7 HUF/L sur boissons sucrées ; 200 HUF/L sur concentrés de sirop

Notes : Ce tableau résume la conception des taxes sur les boissons sucrées en Europe. Source : UNC, Carolina Population Center.

La justification économique de ces taxes repose sur le concept pigouvien de taxes correctrices des externalités, c'est-à-dire qui corrigent les inefficacités du marché en internalisant les externalités négatives et ainsi en alignant le coût privé sur le véritable coût social. Leur objectif est de réduire la demande en augmentant les prix tout en générant des recettes fiscales pour le gouvernement. En pratique, les fabricants supportent le coût supplémentaire de la production et décident ensuite de la part de la

⁸La France a introduit sa taxe sur le sucre en 2012, qui appliquait initialement un taux de 0,07 euros par litre sur les boissons sucrées contenant des sucres ajoutés ou des édulcorants ; elle a été révisée en 2018 pour taxer les boissons en fonction de leur teneur en sucre ajouté, selon une approche progressive. Au Royaume-Uni, la Soft Drinks Industry Levy, mise en place en 2018, taxe les boissons contenant plus de 5 grammes de sucre pour 100 millilitres à 18 pence par litre et celles contenant plus de 8 grammes par 100 millilitres à 24 pence par litre.

taxe qui sera répercutée sur les consommateurs.⁹ Anticiper la répercussion de la taxe n'est pas simple car les entreprises peuvent choisir soit d'absorber une partie de la taxe (le prix final augmente de moins que le montant de la taxe), soit de répercuter entièrement la taxe (le prix final augmente du montant de la taxe), soit d'augmenter leurs marges (le prix final augmente de plus que le montant de la taxe). Les consommateurs ajusteront ensuite leur consommation en fonction de l'augmentation du prix, selon leur sensibilité aux variations de prix (élasticité-prix de la demande). La théorie économique prédit que si le prix des produits taxés augmente, la consommation diminuera, incitant les consommateurs à se tourner vers des alternatives relativement moins chères et potentiellement plus saines, qui ne sont pas soumises à la taxe. Cela peut soit dissuader la consommation d'une famille de produits entière, soit encourager des substitutions vers des produits plus sains au sein de la même famille.

Cependant, la justification de l'intervention publique, en particulier à travers les "taxes sur les comportements à risque", est complexe [Allcott et al., 2019b; Bhattacharya and Sood, 2011]. D'une part, ces taxes peuvent améliorer la santé publique et générer des recettes fiscales pour le gouvernement. D'autre part, elles sont souvent critiquées pour leur caractère paternaliste et régressif, car ces produits sont consommés de manière disproportionnée par les personnes à faibles revenus. L'efficacité de telles taxes dépend d'un design minutieux, notamment en ciblant les bons produits et consommateurs [Allcott et al., 2019a; Dubois et al., 2020; Griffith et al., 2019] et en tenant compte du pouvoir de marché [O'Connell and Smith, 2024]. L'évaluation des politiques nutritionnelles peut être fondée sur plusieurs critères, tels que la consommation de sucre ajouté, le statut d'obésité, les recettes fiscales ou les impacts sur le bien-être.¹⁰

Dans le Chapitre 3, nous développons un nouveau cadre pour évaluer l'optimalité de la conception fiscale en termes de bien-être global (en intégrant le coût social du

⁹Les fabricants peuvent également réagir en modifiant la composition du produit. Toutefois, la littérature se concentre principalement sur les changements de prix plutôt que sur les changements de qualité à la suite de l'introduction d'une taxe (voir Allais et al. [2024] pour un exemple sur la reformulation des recettes des produits).

¹⁰Dans un rapport commandé pour le projet Science and Technology in childhood Obesity Policy, nous avons évalué les impacts potentiels de nouvelles politiques fiscales sur le sucre ajouté en Europe [disponible ici]. Nous avons évalué divers designs fiscaux hypothétiques et leurs effets sur les achats de boissons non alcoolisées et de biscuits, ainsi que sur l'apport en sucre provenant de ces produits, notamment chez les enfants et les adolescents. Nous avons analysé des données issues d'achats de ménages pour proposer des évaluations ex-ante et des comparaisons des effets de différents designs fiscaux au Royaume-Uni, en France et en Espagne. Nos résultats ont souligné l'importance de cibler les catégories alimentaires mauvaises pour la santé ayant une forte sensibilité au prix et de fixer les niveaux de taxe de manière appropriée en fonction de la distribution des achats par rapport au contenu en sucre [Allais et al., 2022].

sucre).

4 Méthodologie de la Thèse

Cette thèse, basée sur une approche empirique, s'appuie sur un ensemble riche de données (données administratives des entreprises et des individus, enquêtes et données d'achats des ménages) ainsi que sur diverses méthodes.

4.1 Sources de données

Cohorte épidémiologique. Les chapitres 1 et 2 exploitent la cohorte épidémiologique *Constances*, un panel d'environ 450 000 individus représentatifs des salariés du secteur privé français âgés de 18 à 69 ans à l'inclusion (voir [Zins et al. \[2015\]](#)). Les consommations de soins des individus est fournie par la Caisse Nationale d'Assurance Maladie avec les dossiers administratifs de remboursement pour les visites médicales, les hospitalisations, les médicaments délivrés, les arrêts maladie et les affections de longue durée (de 2007 à 2019). Les historiques d'emploi sont fournis par la Caisse Nationale d'Assurance Vieillesse et comprennent le statut d'emploi, les revenus, la catégorie socio-professionnelle et l'identifiant du lieu de travail (de 1950 à 2019). Les données individuelles sont appariées aux données administratives des entreprises grâce à l'identifiant du lieu de travail. Des indicateurs contextuels locaux sont ajoutés en fonction du lieu de résidence au niveau de la *commune* ou du *département* (taux de chômage local, mesure de la tension sur le marché du travail et offre de soins de santé locale).

Sources d'entreprises. Les chapitres 1 et 2 utilisent le jeu de données apparié employeur-employé (Déclarations Annuelles des Données Sociales, DADS - Postes), qui fournit des informations détaillées sur les périodes d'emploi pour chaque travailleur, établissement et année, renseignés dans les déclarations fiscales obligatoires (2002-2019). Ce jeu de données est utilisé pour collecter les caractéristiques des entreprises (telles que le nombre d'employés et la structure professionnelle) et pour identifier les fermetures d'établissements. Le chapitre 2 utilise les données administratives des bilans financiers (FICUS-FARE et Bénéfices Réels Normaux, BRN), qui incluent des données financières et des bilans des entreprises françaises, collectées dans les déclarations fis-

cales obligatoires (2013-2019). Ces données sont utilisées pour extraire les caractéristiques des entreprises (telles que le secteur, le chiffre d'affaires et la valeur ajoutée) et pour calculer les investissements dans l'équipement industriel et les machines. Le chapitre 2 exploite également les bases de données des douanes françaises, qui fournissent des informations détaillées au niveau de la transaction sur les valeurs et quantités d'importations par produit et pays d'origine, désagrégées par entreprise et année (1993-2019). Les codes produits détaillés à 8 chiffres de la Nomenclature Combinée (CN-8) sont utilisés pour identifier les importations de biens d'automatisation et de robotique.

Données d'achats alimentaires. Le chapitre 3 utilise un panel de consommateurs représentatif provenant de Kantar Worldpanel au Royaume-Uni (2017). Le jeu de données inclut les enregistrements quotidiens d'achats alimentaires de près de 24 500 ménages. Il fournit des informations détaillées sur les ménages (telles que le statut socio-économique et l'âge, le poids, la taille de chaque membre) et sur les caractéristiques des produits (telles que la marque, la taille, la quantité, le prix, la teneur en sucre).

4.2 Méthodes empiriques

Différence-en-différences. Les chapitres 1 et 2 s'appuient sur des designs quasi-expérimentaux pour estimer les effets causaux d'événements de carrière sur la santé des travailleurs. Dans le chapitre 1, j'exploite les fermetures d'établissements pour identifier les pertes d'emplois exogènes. Dans le chapitre 2, je détecte l'adoption des technologies d'automatisation sur le lieu de travail en utilisant i) des variations dans le stock d'équipement industriel et de machines, et ii) des importations de biens d'automatisation et de robots industriels. Dans les deux chapitres, j'applique une stratégie de différence-en-différences (DiD), en comparant les changements dans les indicateurs au fil du temps entre un groupe traité et un groupe de contrôle soigneusement choisi. Cette méthode permet d'isoler l'impact de l'événement en mesurant la différence des indicateurs avant et après pour les deux groupes et en éliminant les effets temporels ou les facteurs confondants qui auraient pu influencer les deux groupes de manière similaire. L'hypothèse clé sous-jacente à cette approche est l'hypothèse des tendances parallèles, selon laque-

lle, en l'absence de l'événement, les groupes traité et de contrôle auraient suivi des trajectoires similaires. J'intègre les avancées récentes de la littérature DiD [[de Chaisemartin and D'Haultfoeuille, 2020a](#); [Goodman-Bacon, 2021](#)] et applique des estimateurs alternatifs [[Callaway and Sant'Anna, 2021](#); [Cengiz et al., 2019](#); [de Chaisemartin and D'Haultfoeuille, 2020b](#); [Sun and Abraham, 2021](#)] pour garantir la robustesse des résultats.

Évaluation ex-ante & estimation structurelle. La méthodologie du chapitre 3 repose sur le cadre d'évaluation contrefactuelle de l'organisation industrielle structurelle, tel qu'introduit par les travaux fondateurs de [Berry et al. \[1995, 2004\]](#). Cette analyse est enrichie par l'incorporation d'un cadre réglementaire où le bien-être total est composé du surplus des consommateurs, des profits des entreprises, des recettes fiscales et des coûts externes de la consommation excessive de sucre. Ces externalités incluent les coûts de santé associés à des affections telles que la prise de poids, le diabète de type 2 et les maladies cardiovasculaires [[Allcott et al., 2019b](#)]. En prenant en compte ces coûts, l'analyse du bien-être capture les effets potentiellement divergents sur les consommateurs, les entreprises, les recettes fiscales et la santé publique. Pour estimer les quatre composantes du bien-être, nous employons une approche économétrique structurelle en trois étapes, suivant [Bonnet and Réquillart \[2013\]](#). Premièrement, nous estimons un modèle de demande à choix discret des achats alimentaires (modèle logit à coefficients aléatoires) qui capture les substitutions à l'intérieur et entre les variétés de boissons non alcoolisées, identifiant les préférences spécifiques des ménages et les courbes de demande. Deuxièmement, nous modélisons l'offre comme un oligopole avec des produits différenciés, où les entreprises se font concurrence dans un cadre de Bertrand-Nash, similaire aux méthodes utilisées dans [Berry et al. \[1995\]](#) et [Nevo \[2001\]](#). Cette étape nous permet d'estimer les marges prix-coût et les coûts de production pour la répercussion de la taxe. Enfin, nous simulons des scénarios contrefactuels pour évaluer les impacts sur le bien-être des interventions réglementaires.

5 Résumé des Chapitres et Contributions

5.1 Chapitre 1 - Perte d'Emploi et Indicateurs de Santé : le cas de la France

Motivation. Le premier chapitre explore les effets immédiats et à long terme de la perte d'emploi sur la santé. Dans de nombreux pays, les travailleurs sont protégés contre la perte de revenu due au chômage, mais l'impact potentiel sur la santé reçoit une attention limitée. Ceci est surprenant, étant donné les changements significatifs qui accompagnent la perte d'emploi. La réduction des revenus peut entraîner du stress lié à l'insécurité économique et, dans certains cas, conduire au renoncement aux soins. De plus, la perte d'emploi est souvent associée à des changements de mode de vie, tels qu'un plus grand isolement et la perte d'un emploi du temps structuré. En comparant l'état de santé des personnes au chômage et des travailleurs en emploi, on observe que les chômeurs sont plus susceptibles d'être en mauvaise santé (physique et mentale). Par exemple, ils ont 48% de chances en plus d'utiliser des antidépresseurs et des anxiolytiques. Cela soulève la question de savoir si ces problèmes de santé sont une cause ou une conséquence de la perte d'emploi. D'une part, une détérioration de la santé peut mener au chômage ; d'autre part, l'événement même de la perte d'emploi pourrait contribuer à une mauvaise santé. Ce chapitre examine si la perte d'emploi engendre des coûts de santé cachés, en se concentrant sur la France, où les allocations chômage sont généreuses et l'assurance maladie est universelle.

Données & Méthodes. J'utilise la cohorte épidémiologique française *Constances*, qui fournit un ensemble de données unique au niveau individuel à partir de plusieurs sources d'information. Elle comprend des données administratives sur l'historique du marché du travail (telles que l'emploi, les revenus et l'identifiant du lieu de travail), la consommation de soins (telles que les visites chez le médecin, les séjours hospitaliers et les médicaments délivrés), et des données d'enquête sur la santé subjective et les comportements à risque. Le principal défi empirique pour répondre à cette question est d'isoler les pertes d'emplois qui ne sont pas dues à des problèmes de santé préexistants. Il est crucial d'exclure les individus qui ont perdu leur emploi à cause de problèmes de santé antérieurs. Ainsi, l'objectif est de trouver un choc qui entraîne la perte

d'emploi et qui soit indépendant de la santé des travailleurs. Pour ce faire, j'utilise la stratégie d'identification proposée dans l'article fondateur de [Jacobson et al. \[1993\]](#), qui examine les licenciements massifs et les fermetures d'établissements. Je repère ces événements à l'aide de bases de données administratives employeur-employé appariées. Dans le cas des fermetures d'établissements, tous les travailleurs sont licenciés, indépendamment de leur état de santé, ce qui permet d'identifier un groupe de travailleurs ayant subi une perte d'emploi involontaire, appelés « travailleurs traités ». Je les compare à un groupe de contrôle de travailleurs partageant des caractéristiques similaires, mais n'ayant pas subi de perte d'emploi. Plus précisément, je compare l'évolution des indicateurs de santé entre les deux groupes à l'aide d'une méthode de différence-en-différences. L'hypothèse clé est qu'en l'absence de perte d'emploi, les trajectoires de santé des deux groupes auraient été similaires. Cela permet d'estimer les indicateurs de santé contrefactuels pour le groupe traité, en montrant ce qu'aurait été leur santé si la perte d'emploi ne s'était pas produite.

Résultats & Contributions. Les résultats révèlent d'importants impacts économiques et sanitaires à long terme de la perte d'emploi. Les travailleurs licenciés connaissent une forte baisse de leurs revenus annuels, perdant 5 000 € la première année, avec des revenus encore 3 000 € inférieurs cinq ans plus tard. La santé mentale est particulièrement affectée, avec une augmentation de l'utilisation d'antidépresseurs et d'anxiolytiques de 19% et une hausse de 32% de l'utilisation de somnifères, tant à court qu'à long terme. Les travailleurs licenciés sont également plus enclins à consulter des psychiatres et à être hospitalisés pour des problèmes de santé mentale peu après la perte d'emploi. La santé physique se détériore également, avec une augmentation des visites chez le médecin, des admissions à l'hôpital, et une probabilité plus de deux fois supérieure de recevoir des pensions d'invalidité. Ces effets sont particulièrement importants pour des groupes vulnérables, comprenant les travailleurs plus âgés, les travailleurs peu qualifiés et ceux vivant dans des régions avec taux de chômage élevé. Je montre que le mécanisme des nouvelles opportunités d'emploi joue un rôle crucial, car l'augmentation des consommations des soins de santé mentale est concentrée chez les travailleurs à risque élevé, c'est-à-dire ceux ayant un risque plus élevé de chô-

mage. De plus, la détérioration de la santé mentale est principalement causée par le choc immédiat de la perte d'emploi, plutôt que par la durée du chômage, l'utilisation d'antidépresseurs augmentant juste après le licenciement.

5.2 Chapitre 2 - Automatisation sur le Lieu du Travail: Implications pour la Santé des Travailleurs

Ce chapitre est co-écrit avec Pauline Lesterquy (Banque de France, CREST).

Motivation. Au cours des derniers siècles, les historiens ont documenté des révolutions industrielles successives, chaque vague de progrès technologique ayant profondément modifié l'emploi et le contenu des tâches professionnelles. Ce chapitre se concentre sur les deux dernières vagues, marquées par l'avènement de l'automatisation et de la robotique. En particulier, nous cherchons à comprendre l'impact de l'adoption de ces innovations au travail sur la santé des travailleurs. Une large littérature a documenté les effets directs de ces avancées technologiques sur la nature du travail et les tâches accomplies par les travailleurs. La perspective négative met en avant le remplacement du travail humain par des machines automatisées, entraînant des pertes d'emplois [Acemoglu and Restrepo, 2020; Autor et al., 2003; Bessen et al., 2023; Goos, 2018]. En revanche, la vision positive souligne les gains de productivité générés par le changement technologique [Aghion et al., 2024; Bonfiglioli et al., 2024; Graetz and Michaels, 2018]. Ces transformations des processus de production peuvent également influencer le bien-être des travailleurs. Cependant, l'impact global sur la santé des travailleurs reste peu étudié. Les effets attendus sont ambigus. Sur le plan de la santé mentale, le risque de déplacement pourrait induire du stress, mais l'introduction de tâches plus valorisantes pourrait potentiellement accroître la satisfaction au travail. En ce qui concerne la santé physique, des conditions de travail améliorées, comme des tâches moins physiquement exigeantes, pourraient bénéficier aux travailleurs, bien que le risque de blessures au travail en raison de la présence de machines pourrait augmenter. De plus, les changements dans l'intensité du travail pourraient affecter tant la santé mentale que physique. Ce chapitre quantifie empiriquement les effets moyens et examine également les sous-groupes qui pourraient être affectés différemment.

Données & Méthodes. Nous nous concentrons sur le secteur manufacturier en France entre 2007 et 2019. Pour le premier ensemble d’analyses au niveau de l’entreprise, nous utilisons plusieurs ensembles de données administratives sur les entreprises (douanes, bilans et base de données employeur-employé appariées) pour capturer l’adoption d’équipements industriels et de technologies d’automatisation par les entreprises. Puisque l’observation directe de l’adoption de l’automatisation est difficile, nous utilisons deux proxies clés. Le premier proxy repose sur les données de bilans comptables concernant les équipements industriels et les machines pour estimer les investissements des entreprises dans le capital physique. Le second proxy se concentre sur les importations de biens d’automatisation (ou alternativement de robots industriels) au niveau des entreprises, tirées des données douanières détaillées. Nous détectons les événements d’automatisation en considérant les investissements importants dans les biens de capital manufacturier à l’aide de ces deux proxies. Nous cherchons à capturer les changements dans les processus de production résultant de l’adoption de l’automatisation. Pour le deuxième ensemble d’analyses, au niveau des travailleurs, nous exploitons les données administratives de consommations de soins individuelles issus de la cohorte épidémiologique française *Constances*. Ces données sont appariées avec les ensembles de données des entreprises grâce à l’identifiant du lieu de travail. Cela nous permet d’identifier les travailleurs exposés aux changements technologiques au sein des entreprises. L’ensemble de données riche sur les consommations de soins nous permet d’étudier les effets de l’automatisation sur la santé mentale et physique des travailleurs et d’examiner les mécanismes sous-jacents. Nous nous appuyons également sur une méthodologie de différence-en-différences, où nous comparons un groupe d’entreprises (ou de travailleurs) traitées à un groupe d’entreprises (ou de travailleurs) témoins, l’hypothèse d’identification étant que leurs indicateurs auraient évolué de la même manière en l’absence de l’adoption de la technologie d’automatisation.

Résultats & Contributions. Premièrement, nous caractérisons ce qui se passe lorsqu’une entreprise amorce un changement technologique. Nous observons que les entreprises augmentent le nombre d’employés avec des effets de composition, car la part d’ouvriers augmente. Cela entraîne une diminution du salaire horaire moyen. Nous constatons

également une augmentation de la productivité du travail et une diminution de la part du travail dans la valeur ajoutée. Les résultats soutiennent la vision positive de l'automatisation, ce qui signifie que, dans notre contexte, les mécanismes liés à la productivité l'emportent sur ceux liés à la destructions des emplois au niveau de l'entreprise en moyenne. Deuxièmement, nous examinons les conséquences pour les travailleurs. En moyenne, les travailleurs déjà en poste ne subissent pas de changements significatifs dans leurs indicateurs d'emploi (revenus et statut d'emploi), mais les travailleurs plus âgés font face à un risque élevé de perte d'emploi. Nous documentons une détérioration de la santé mentale, avec une augmentation moyenne de 27% de l'utilisation des antidépresseurs et des anxiolytiques. La santé physique reste globalement inchangée, à l'exception d'une diminution de l'utilisation de médicaments anti-inflammatoires chez les travailleurs à bas salaire. En outre, les travailleurs plus âgés et les ouvriers présentent une hausse marquée des pensions d'invalidité, et les travailleurs plus âgés prennent en moyenne six jours supplémentaires d'arrêt maladie.

5.3 Chapitre 3 - Les Taxes sur le Sucre sont-elles Bien Conçues ? Etude empirique sur le marché des boissons sans alcool au Royaume-Uni

Ce chapitre est co-écrit avec Olivier Allais (INRAE, PSAE), Céline Bonnet (INRAE, TSE) et Maxime Tranchard (INRAE, PSAE).

Motivation. Les décideurs politiques se tournent de plus en plus vers les taxes nutritionnelles comme outil pour orienter le comportement des consommateurs vers des régimes alimentaires plus sains. À ce jour, plus de 50 pays ont mis en place des taxes sur les boissons sucrées. Cependant, l'efficacité de ces taxes dépend largement de leur conception. La recherche théorique et empirique suggère que les taxes devraient idéalement être proportionnelles au degré de préjudice causé par la consommation. Cependant, dans la pratique, aucune des taxes nutritionnelles mises en œuvre jusqu'à présent ne suit une structure strictement linéaire (taux de taxe proportionnel à la teneur en sucre). En effet, la plupart d'entre elles présentent des conceptions par paliers avec des seuils qui varient dans leur proximité. Les mises en œuvre réelles peuvent donc différer de la théorie économique en raison de contraintes de faisabilité. Cela soulève

la question suivante : ces taxes par paliers peuvent-elles être optimales et, si oui, comment leurs principaux composants doivent-ils être définis ?

Données & Méthodes. Nous utilisons des données d'achats de ménages sur le marché des boissons non alcoolisées au Royaume-Uni en 2017. Ce panel représentatif fournit des informations détaillées sur les caractéristiques des ménages et des produits, notamment la marque, la taille, la quantité, le prix et la teneur en sucre. Nous développons un cadre réglementaire dans lequel la fonction objective est le bien-être total, englobant le surplus du consommateur, le profit des entreprises, les recettes fiscales et le coût social du sucre (représenté par les externalités sanitaires dues à la consommation excessive de sucre). L'objectif du régulateur est de sélectionner la conception de la taxe qui maximise le bien-être total en fonction de trois paramètres : le nombre de seuils, l'emplacement des seuils (en g/100mL), et les taux de taxe (en euro cents/L). Ce cadre flexible permet d'explorer un large éventail de structures fiscales, des taxes proportionnelles à la teneur en sucre aux conceptions par paliers avec un nombre variable de seuils. Nous nous appuyons sur l'estimation du coût social du sucre réalisée par [Rischbieth et al. \[2020\]](#), à environ 27 € par kilogramme de consommation de sucre. Nous évaluons les effets des politiques fiscales nutritionnelles en utilisant des méthodes économétriques structurelles ex-ante qui capturent à la fois les réactions des consommateurs et des entreprises aux changements de taxe. En suivant la méthodologie en trois étapes de [Bonnet and Réquillart \[2013\]](#), nous estimons d'abord la demande à l'aide d'un modèle logit à coefficients aléatoires pour comprendre les préférences des consommateurs en matière de prix et de qualité. Ensuite, nous modélisons l'offre comme une concurrence oligopolistique, ce qui nous permet d'analyser les stratégies de prix, les coûts marginaux et le passage des taxes par les entreprises. Enfin, nous menons des simulations contrefactuelles de différents scénarios de taxation pour étudier leurs effets sur le bien-être total.

Résultats & Contributions. Nos résultats indiquent que les structures fiscales par paliers, qui prennent en compte les réponses stratégiques des entreprises, peuvent générer des gains de bien-être bien plus importants que la taxe actuelle sur les bois-

sons sucrées du Royaume-Uni (Soft Drinks Industry Levy - SDIL). En particulier, la conception optimale impose des taxes plus élevées sur les produits riches en sucre, encourageant les entreprises à baisser les prix des alternatives à faible teneur en sucre. Cela conduit non seulement à une amélioration des indicateurs en matière de santé publique, mais aussi à une augmentation du surplus des consommateurs tout en maintenant la rentabilité des entreprises. Ces résultats suggèrent que les décideurs politiques devraient soigneusement prendre en compte le comportement stratégique des entreprises lors de la conception des politiques de taxation du sucre afin de maximiser les bénéfices en termes de bien-être.



Bibliography of the Introduction

Acemoglu, Daron and David Autor, “Tasks and Technologies: Implications for Employment and Earnings,” *Handbook of Labor Economics*, 2011, 4, 1043–1171.

– **and Pascual Restrepo**, “Automation and New Tasks: How Technology Displaces and Reinstates Labor,” *Journal of Economic Perspectives*, 2019, 33 (2), 3–30.

– **and –**, “Robots and Jobs: Evidence from US Labor Markets,” *Journal of Political Economy*, 2020, 128 (6), 2553–2584.

– **, Claire Lelarge, and Pascual Restrepo**, “Competing with Robots: Firm-Level Evidence from France,” *AEA Papers and Proceedings*, 5 2020, 110, 383–388.

Aghion, Philippe, Céline Antonin, Simon Bunel, and Xavier Jaravel, “Modern Manufacturing Capital, Labor Demand, and Product Market Dynamics: Evidence from France,” *Working paper*, 2024.

Allais, Olivier, Céline Bonnet, Pauline Leveneur, and Maxime Tranchard, “Assessment of the potential impacts of new fiscal and regulatory policies on added sugar in Europe,” *Deliverable - Science & Technology in childhood Obesity Policy*, 2022.

– **, – , Vincent Réquillart, Marine Spiteri, and Maxime Tranchard**, “Sugar-based tax design for healthier consumption: Empirical evidence in the French Dessert market,” *Working Paper*, 2024.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, “Regressive Sin Taxes, with an Application to the Optimal Soda Tax,” *The Quarterly Journal of Economics*, 05 2019, 134 (3), 1557–1626.

– **, – , and –**, “Should We Tax Sugar-Sweetened Beverages? An Overview of Theory and Evidence,” *Journal of Economic Perspectives*, 2019, 33 (3), 202–227.

Autor, David, “Why are there still so many jobs? the history and future of workplace automation,” *Journal of Economic Perspectives*, 2015, 29, 3–30.

– **and David Dorn**, “The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market,” *American Economic Review*, 2013, 103, 1553–1597.

– **, Frank Levy, and Richard Murnane**, “The Skill Content of Recent Technological

-
- Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, 118, 1279–1333.
- Berniell, Inés and Jan Bietenbeck**, "The Effect of Working Hours on Health," *Economics & Human Biology*, 2020, 39 (100901).
- Berry, Steven, James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, 63 (4), 841–890.
- , –, and –, "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of political Economy*, 2004, 112 (1), 68–105.
- Bessen, James, Maarten Goos, Anna Salomons, and Wiljan van den Berge**, "What Happens to Workers at Firms that Automate?," *The Review of Economics and Statistics* (forthcoming), 2023.
- Bhattacharya, Jay and Neeraj Sood**, "Who Pays for Obesity?," *Journal of Economic Perspectives*, 2011, 2 (25), 139–158.
- Blasco, Sylvie, Julie Rochut, and Bénédicte Rouland**, "Displaced or Depressed? The Effect of Working in Automatable Jobs on Mental Health," *IZA Discussion Paper Series*, 2022.
- Bonfiglioli, Alessandra, Rosario Crinò, Harald Fadinger, and Gino Gancia**, "Robot Imports and Firm-Level Outcomes," *Economic Journal* (forthcoming), 2024.
- Bonnet, Céline and Vincent Réquillart**, "Tax incidence with strategic firms in the soft drink market," *Journal of Public Economics*, 2013, 106, 77–88.
- Brynjolfsson, Erik and Andrew McAfee**, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: Norton, 2014.
- Burgard, Sarah A. and Katherine Lin**, "The health benefits of employment," *Journal of Health and Social Behavior*, 2013, 54 (2), 251–265.
- Callaway, Brantly and Pedro Sant'Anna**, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Capacci, Sara, Olivier Allais, Celine Bonnet, and Mario Mazzocchi**, "The impact of the French soda tax on prices and purchases. An ex post evaluation," *forthcoming in*

- PlosOne*, 2019.
- Caroli, Eve and Mathilde Godard**, “Does job insecurity deteriorate health?,” *Health Economics*, 2016, 25, 131–147.
- Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer**, “The effect of minimum wages on low-wage jobs,” *The Quarterly Journal of Economics*, 2019, 134 (3), 1405–1454.
- DARES**, “Quelles étaient les conditions de travail en 2019, avant la crise sanitaire ?,” *DARES Analyses*, 2021, 44.
- de Chaisemartin, Clément and Xavier D’Haultfoeuille**, “Two-way fixed effects estimators with heterogeneous treatment effects,” *American Economic Review*, 2020, 110 (9), 2964–2996.
- **and –**, “Difference-in-Differences Estimators of Intertemporal Treatment Effects,” *Working paper*, 2020.
- Dubois, Pierre, Paulo Albuquerque, Olivier Allais, Céline Bonnet, Patrice Bertail, Pierre Combris, Saadi Lahlou, Natalie Rigal, Bernard Ruffieux, and Pierre Chandon**, “Effects of front-of-pack labels on the nutritional quality of supermarket food purchases: evidence from a large-scale randomized controlled trial,” *Journal of the Academy of Marketing Science*, 2021, 49, 119–138.
- **, Rachel Griffith, and Martin O’Connell**, “How well targeted are soda taxes?,” *American Economic Review*, 2020, 110 (11), 3661–3704.
- European Observatory on Health Systems and Policies**, “France: Country Health Profile 2023,” 2023. Accessed: 2024-09-24.
- Frey, Carl and Michael Osborne**, “The future of employment: How susceptible are jobs to computerisation?,” *Technological Forecasting and Social Change*, 2017, 114, 254–280.
- Goldin, Claudia and Lawrence F. Katz**, “The Origins of Technology-Skill Complementarity,” *The Quarterly Journal of Economics*, 1998, 113 (3), 693–732.
- Goodman-Bacon, Andrew**, “Difference-in-differences with variation in treatment timing,” *Journal of Econometrics*, 2021, 225, 254–277.

-
- Goos, Maarten**, “The impact of technological progress on labour markets: Policy challenges,” *Oxford Review of Economic Policy*, 7 2018, 34, 362–375.
- **and Alan Manning**, “Lousy and Lovely Jobs: The Rising Polarization of Work in Britain,” *The Review of Economics and Statistics*, 2007, 89, 118–133.
- Graetz, Georg and Guy Michaels**, “Robots at work,” *The Review of Economics and Statistics*, 12 2018, 100, 753–768.
- Griffith, Rachel, Martin O’Connell, and Kate Smith**, “Tax design in the alcohol market,” *Journal of Public Economics*, 2019, 172, 20–35.
- Humlum, Anders**, “Robot Adoption and Labour Market Dynamics,” *Working paper*, 2021.
- INSERM**, “Mental health in France: Current status and challenges,” 2020.
- Jacobson, Louis, Robert LaLonde, and Daniel Sullivan**, “Earnings Losses of Displaced Workers,” *American Economic Review*, 1993, 83 (4), 685–709.
- Lobstein, Tim and Margot Neveux**, “A review of systematic reviews of the impact on children of three population-wide policies,” *Deliverable - Science & Technology in childhood Obesity Policy*, 2021.
- Marmot, Michael and Stephen Stansfeld**, “Work and health: The role of employment and working conditions,” *The Lancet*, 2003, 362 (9386), 1001–1006.
- Nevo, Aviv**, “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 2001, 69 (2), 307–342.
- Obesity Evidence Hub**, “Countries that have taxes on sugar-sweetened beverages (SSBs),” Technical Report 2022. Available online at <https://www.obesityevidencehub.org.au/collections/prevention/countries-that-have-implemented-taxes-on-sugar-sweetened-beverages-ssbs>, accessed March, 2024.
- O’Connell, Martin and Kate Smith**, “Optimal Sin Taxation and Market Power,” *American Economic Journal: Applied Economics*, 2024, 16 (4), 34—70.
- Public Health England**, “Results of the National Diet and Nutrition Survey (NDNS) rolling programme for 2014 to 2015 and 2015 to 2016,” Technical Report 2018.

Available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/699241/NDNS_results_years_7_and_8.pdf, accessed March, 2019.

Rischbieth, Amanda, George Serafeim, and Katie Trinh, “Accounting for product impact in the consumer-packaged foods industry,” *Harvard Business School Working Paper*, 2020, 21-051.

Santé Publique France, “Chronic diseases in France: Prevalence and impact,” 2021.

Sullivan, Daniel and Till von Wachter, “Job Displacement and Mortality: An Analysis using administrative Data,” *Quarterly Journal of Economics*, 2009, 124 (2), 1265–1306.

Sun, Liyang and Sarah Abraham, “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects,” *Journal of Econometrics*, 2021, 225 (2), 175–199.

World Bank, “Digital Dividends,” *World Development Report 2016*, 2016.

Yakymovych, Yaroslav, “Consequences of job loss for routine workers,” *IFAU Working Paper*, 2022, 15.

Zins, Marie, Marcel Goldberg, and Constances-team, “The French CONSTANCES population-based cohort: design, inclusion and follow-up,” *European Journal of Epidemiology*, 2015, 30 (12), 1317–1328.



Chapter 1

Job Loss and Health Outcomes: Evidence from France

Abstract

This paper investigates the consequences of job loss on a wide range of health outcomes in France, a country with generous unemployment benefits and universal health insurance. I build a new dataset by combining administrative panel data from a French epidemiological cohort with matched employer-employee data. Focusing on exogenous job losses resulting from establishment closures, I implement a staggered difference-in-differences approach in which I examine the evolution of the health outcomes of 20,000 treated workers relative to those of a comparable control group. My results highlight the adverse effects of job loss on mental health. Displaced workers experience a 19% increase in the use of antidepressants and anxiolytics and a 20% rise in the likelihood of consulting a psychiatrist in the short run. Hospital admissions increase by 13%. Additionally, the probability of receiving disability benefits more than doubles after displacement. Heterogeneity analyses reveal that specific vulnerable populations, including older workers, low-skilled workers, and workers living in areas with high unemployment rates, are more severely impacted by job loss. The perspective of re-employment also plays an important role in mitigating the consequences on mental health.

Keywords: Job Loss, Plant closure, Physical and mental health, Health insurance

JEL classification: I12, I14, J65

1 Introduction

Job loss has persistent adverse effects on employment [[Couch and Placzek, 2010](#); [Jacobson et al., 1993](#); [Schmieder et al., 2022](#)], displaying varying magnitudes across countries [[Bertheau et al., 2022](#)] and individuals [[Athey et al., 2024](#)]. Unemployment insurance generally addresses the risk of income reduction and is designed to balance the insurance value and the moral hazard of new job seekers. Yet, the policy design disregards the potential adverse health effects stemming from job loss despite evidence showing a link between poor (mental) health status and unemployment [[OECD, 2019](#)]. After job loss, the decrease in income can induce stress from economic insecurity, while lifestyle changes such as isolation, loss of self-esteem, and the absence of a structured schedule can impede mental health.¹ If job loss indeed is the cause of adverse health status, then public policies could greatly benefit from a more comprehensive understanding of the connection between unemployment and health.

This paper evaluates the impact of job loss on health in France, where a comprehensive social security system provides generous unemployment insurance with active labor market policies and universal health coverage.² I am investigating whether job loss is detrimental to physical and mental health within a context where individuals are well-insured. This paper contributes to the existing literature on several dimensions. First, I leverage a very detailed administrative healthcare dataset which allows me to provide an extensive understanding of the impact of job loss on a comprehensive set of health indicators. Importantly, I can compare the effect of job loss on this set of health indicators within the same study design, whereas other papers often focus on only one healthcare dimension. Second, I estimate the dynamic effects of job loss following the recent developments in the staggered difference-in-differences literature. Robustness checks, including alternative specifications and estimators, validate the consistency of the results.³ Finally, I conduct new heterogeneity analyses to identify vulnerable pop-

¹In an experiment, [Hussam et al. \[2022\]](#) highlights that the psychosocial value of employment goes beyond its role as an income source. Being employed enhances psychosocial well-being significantly more than receiving cash alone.

²France can be considered as an intermediate case between the United States, with its minimal safety net, and the Nordic countries, which offer extensive health and unemployment insurance. France is closer to the Nordic countries but stands out for its higher share of long-term unemployment.

³I run a series of robustness checks. First, I use the standard specification of [Jacobson et al. \[1993\]](#) and [Sullivan and von Wachter \[2009\]](#). Second, I use a different control group (the not-yet-treated individuals). Finally, I implement the estimators developed

ulations that bear a higher cost from job loss, encompassing demographics, job characteristics, and local contextual factors. Additionally, I explore some of the underlying mechanisms driving the effects on mental health.

I create a new panel by matching several survey and administrative data, allowing me to study the impact of exogenous job loss on health and economic outcomes in France between 2007 and 2019. First, I use a representative panel of private sector employees, sampled for the *Constances* epidemiological cohort, which provides detailed administrative information on healthcare and labor market histories. I consider comprehensive physical and mental health indicators based on healthcare records (visits to doctors, hospital stays and prescribed drugs). I also leverage information about individuals' careers, especially the identifier of the establishments where they have worked. Second, I use the French matched employer-employee dataset and I identify establishment closures based on worker flows. Finally, I match the worker-level panel with information on health outcomes and employment history to the establishment-level dataset identifying establishment closures based on the workplace identifier.

Estimating the effect of job loss on health outcomes is challenging, as reasons for job loss are likely to be endogenous to health conditions, potentially leading to non-causal estimates. To address the potential reverse causality and selection bias, I follow the standard approach in the literature by focusing on individuals who lost their jobs due to establishment closures [Couch and Placzek, 2010; Jacobson et al., 1993]. This event can reasonably be considered exogenous to workers' health because all employees lose their jobs regardless of their health conditions. Additionally, I define a comparable control group of never-treated workers using a matching procedure based on employment history and past health conditions. I perform a difference-in-differences event study, with staggered treatment rollout, on approximately 20,000 workers who experienced an establishment closure and 20,000 comparable non-displaced workers (in a similar specification as Britto et al. [2022]; Gerard and Naritomi [2021]; Roulet [2017]; Schmieder et al. [2022]). I estimate the dynamic impact of job loss on health by comparing the evolution of health outcomes of displaced workers with that of a control group

in Callaway and Sant'Anna [2021]; de Chaisemartin and D'Haultfoeuille [2020b]; Sun and Abraham [2021] following the recent literature on difference-in-differences models. The results on health outcomes are similar to the main specification.

with a similar work history prior to job loss. I account for the recent developments in the difference-in-differences literature and my results are robust to the new alternative estimators.⁴ Furthermore, I examine how the effects of job loss may vary, with a focus on vulnerable populations, and the mechanisms related to individual risk in the labor market.

My results indicate that job loss has significant consequences on both economic and health outcomes. Consistent with [Bertheau et al. \[2022\]](#), I find that displaced workers experience a persistent drop in annual earnings (of approximately 20% in the year following job loss and 12% even five years after job loss). This reduction primarily stems from a decreased likelihood of being in employment, leading also to a higher reliance on unemployment benefits. The repercussions of job loss extend beyond economic consequences; my findings also reveal a substantial impact on mental health and the use of health benefits, along with a moderate effect on physical health. The use of antidepressants and anxiolytics increases by around 19% in the short run and the use of sleeping pills increases by 32% up to 8 years after job loss. In the year of displacement, the likelihood of visiting a psychiatrist increases by around 20% and the likelihood of hospital admission increases by 13% with significance in mental health diagnoses and alcohol-related issues. I also show that the probability of receiving disability benefits and extended reimbursements due to long-term conditions more than double. Heterogeneity analyses uncover that specific vulnerable populations, such as older workers, low-skilled workers, and those in areas with high unemployment rates, are more affected by job loss. Gender does not appear as a main driver in heterogeneity. Finally, I show that re-employment opportunities mitigate these effects: only workers with a high risk of staying unemployed increase their use of mental health-related drugs.

Related literature. While job loss exhibits a strong and persistent impact on health in the United States, evidence in Europe is more mixed, suggesting that the social protection system plays an important role. In the United States, job loss is associated with a positive and persistent effect on mortality several years after displacement [[Sullivan](#)

⁴[de Chaisemartin and D'Haultfoeuille \[2020a\]](#); [Goodman-Bacon \[2021\]](#); [Sun and Abraham \[2021\]](#) show the limits of the usual two-way fixed effect models in the case of staggered treatment designs. I implement the alternative estimators developed in [Callaway and Sant'Anna \[2021\]](#); [de Chaisemartin and D'Haultfoeuille \[2020b\]](#); [Sun and Abraham \[2021\]](#).

and von Wachter, 2009] and on mental health [Schaller and Stevens, 2015]. Results are heterogeneous in Europe. Some studies found that job loss increases the risk of mortality in Denmark [Browning and Heinesen, 2012], in the Netherlands [Bloemen et al., 2018] and in Sweden [Eliason and Storrie, 2009a] and is associated with an increase in disability benefits in Norway [Rege et al., 2009]. However, many studies find little effect of job loss on hospitalization for stress-related diseases [Browning et al., 2006] and on mental health, severe physical health outcomes and mortality in Denmark [Roulet, 2017], on the risk of severe cardiovascular diseases and diabetes in Sweden [Bergemann et al., 2019; Eliason and Storrie, 2009b], on health satisfaction, mental health, and hospital visits in Germany [Schmitz, 2011] and on healthcare expenditures in Austria [Kuhn et al., 2009]. In France, Lengagne [2022] finds a positive impact on mortality among older workers but not on disability pension enrollment. Using survey data, Clainche and Lengagne [2019] note an increase in the use of psychotropic drugs among workers who stayed with the firm after a mass layoff.

The paper is organized as follows. Section 2 provides information on the French context and presents the analysis sample and some summary statistics. Section 3 describes the empirical strategy to identify the causal effect of job loss on health outcomes. Section 4 presents the results on the dynamic effects of job loss on health and some heterogeneity analyses. Section 5 discusses the potential mechanisms. Section 6 concludes.

2 Institutional background and data

2.1 Institutional background

Unemployment insurance. France offers a high level of employment protection and a generous unemployment insurance scheme. Eligibility for unemployment benefits requires a minimum of 6 months of work within the past 24 months. The net replacement rate was 68% for a single individual with average previous wage earnings in 2015 (OECD).⁵ The maximum compensation duration is 24 months for individuals un-

⁵Accessed on <https://stats.oecd.org/Index.aspx?DataSetCode=NRR>

der 53, 30 months for those between 53 and 54, and 36 months for individuals aged 54 and above.

Layoffs and plant closures. When an employer intends to close an establishment, a structured process must be followed, starting with informing the establishment's social and economic committee, waiting for the opinion of regional directorates and then noticing the workers.⁶ Therefore, it is essential to account for the possible anticipatory effects. Workers in financially troubled establishments, for instance, might experience health-related issues such as stress before the closure formally takes place.

Health insurance. Health insurance is universal and covers around 65% of health expenditures (except for severe conditions which receive full coverage). To address the remaining costs, individuals have the option to acquire complementary health insurance, either individually or through their employer. Otherwise, the remaining 35% is to be paid by the patients. If this complementary insurance was provided by the employer, job loss could affect healthcare reimbursement. However, individuals are allowed to keep the employer-provided complementary health insurance under certain conditions at the end of a contract.

2.2 Data

2.2.1 Epidemiological cohort

I use data from the *Constances* cohort, a French general-purpose epidemiological panel designed to represent private sector employees aged 18-69 at the time of inclusion (see [Zins et al. \[2015\]](#) for further details). A random sample of individuals was selected based on age, gender, and socio-economic status. Invitations were sent to those selected, resulting in two distinct groups: active participants and the reference cohort. The *active participants* group, consisting of approximately 170,000 individuals, voluntarily agreed to participate. These participants complete self-administered questionnaires covering health, life events, and occupation details (2012-2019), including an

⁶The employer must first consult the company's social and economic committee and pass on the redundancy project to the Dreets. An opinion is then given after 2 months when the number of redundancies is less than 100, 3 months for a number of redundancies between 100 and 249 inclusive and 4 months from 250 redundancies. The employees concerned by the dismissal are summoned and then receive a letter of dismissal. The employer must then inform the Dreets of the economic dismissal. The employee then gives notice of redundancy and the employment contract is terminated following this. The employee can then receive compensation. See <https://entreprendre.service-public.fr/vosdroits/F24648>

initial questionnaire upon inclusion and annual follow-ups.⁷ They are also invited to attend a health examination at one of the 22 Health Screening Centers (HSC).⁸ The *reference cohort*, composed of roughly 300,000 individuals, includes those who were initially sampled but did not respond to the invitation. However, they are still passively monitored through administrative records.⁹ For both active participants and the reference cohort, data on healthcare consumption (2007-2019) and labor market history (1950-2019) are obtained from national administrative databases.

Survey. Participants are invited to complete at-home questionnaires: one at inclusion and then annual follow-ups (2012-2019). The survey data contains information on socio-demographic characteristics, health, behaviors, and occupational factors. I primarily consider questions related to self-reported health, risky behaviors (dietary habits, smoking, and alcohol consumption), and a mental health scale.¹⁰

Labor market history. Administrative data on individuals' complete labor market histories are provided by the National Retirement Insurance Fund ("Caisse Nationale de l'Assurance Vieillesse", CNAV). For each year between 1950 and 2019, the data includes occupation details from employers' reports (*e.g.*, earnings, occupation) and non-work periods reported by social welfare organizations (*e.g.*, unemployment, sick leave, maternity leave, disability benefits). The key variables for analysis are defined as follows. I construct a primary annual employment status (employment, unemployment, sick leave and work accident, disability, maternity leave, retirement, inactivity) based on the number of quarters contributing to old-age pension (see Appendix 1.B for more details). I also compute annual earnings, adjusted to 2002 euros, as the total labor earnings from the private sector, which may include income from multiple employ-

⁷Details on participation in both inclusion and follow-up questionnaires are shown in Figure 1.A.2 in Appendix 1.A. To account for initial non-participation and attrition, adjustment weights and coefficients are computed based on the reference cohort.

⁸Participants are invited to visit an HSC at the time of inclusion and every five years thereafter. They are selected randomly from 19 targeted *départements* associated with a partner Health Screening Center (see map in Appendix 1.A). The representativeness of the sample is preserved, as the demographic, social, and occupational characteristics of the sample closely resemble those of the general French population [Zins et al., 2015].

⁹This cohort was initially used to compute adjustment weights to ensure representativeness of the active participants. We received authorization from CNIL to include them in this study.

¹⁰Subjective health is measured with the question "How do you rate your general health?", which has an 8-level response ranging from "Very good" to "Very bad". Dietary habits are assessed through the question "Do you think your diet is balanced?" with an 8-level response ranging from "Very much" to "Not at all". Smoking behavior is captured by current smoking status and the number of cigarettes smoked per day. Heavy drinking is measured using the minimum average alcohol consumption (in number of drinks) and adherence to daily alcohol consumption recommendations. I also use the 20 items of the CES-D scale [Radloff, 1977], which captures depressive symptoms and includes a synthetic indicator for being in a depressive state. For instance, one item asks "In the past week, how often have you [felt lonely]?", with response options of "Never", "Sometimes", "Often", and "Always".

ers.¹¹ Earnings are set to zero for years in which an individual does not work. Additionally, I am able to link employees to their three main employers (via the unique workplace identifier SIRET) for each year, which is critical for my identification strategy. This administrative labor market history is available for both active participants and the reference cohort.

Healthcare consumption. Administrative data on individuals' healthcare use is provided by the National Health Insurance Fund ("Système National des Données de Santé", SNDS). For each medical act performed between 2007 and 2019, the data includes comprehensive individual-level information from various sources: reimbursement data (*e.g.*, visits to doctors and other health professionals, prescribed drugs, medical devices), records of "long-term conditions" (serious chronic diseases exempt from co-payments), hospital discharge summaries (including diagnoses), and records of medical and technical procedures. I categorize healthcare consumption into three main areas: visits to doctors, hospital stays, and prescribed drugs. For each category, I can distinguish between care related to physical health and mental health. In terms of doctor visits, I differentiate between general practitioners (GPs), mental health specialists (psychiatrists and neuropsychiatrists), and other specialists. Hospital stays are classified according to their main diagnosis, with a focus on cancer, circulatory problems, alcohol-related conditions, and mental health issues. I also account for the use of emergency services. Regarding prescribed drugs, I specifically examine the consumption of antidepressants, anxiolytics, sleeping pills, as well as painkillers, opioids, and anti-inflammatory drugs. The detailed definition of healthcare outcomes is provided in Table 1.C.1, Appendix 1.C. This administrative healthcare data is available for both the active participants and the reference cohort.

Local contextual indicators. I complement individual data from the *Constances* cohort with local contextual indicators. They are matched at the *commune* or *département* level based on the place of residence. To characterize the local labor market, I use the local unemployment rate (from the French National Institute of Statistics) and a synthetic

¹¹This total also includes earnings from apprenticeships, domestic work, au pair or family work, as well as for artists, authors, and entertainment workers. For individuals in "Other Regimes" (*e.g.*, civil servants, local authority or hospital staff), income data is not available, so earnings are left missing for these cases. The data also provides information on old-age pension amounts.

labor market tightness measure.¹² To account for the local healthcare supply, I use an accessibility indicator for General Practitioners.¹³

2.2.2 Establishment closures

Matched employer-employee dataset. I leverage the French matched employer-employee dataset (*Déclarations Annuelles des Données Sociales (DADS) - Postes*) to identify establishment closures. The data are based on employers' mandatory fiscal reports about their employees. A comprehensive dataset is available from 1994 onward. It contains information on jobs at the worker \times establishment \times year level, allowing the tracking of workers' mobility across establishments over time. I identify establishment closures by analyzing the flow of workers between establishments. I also use this dataset to retrieve characteristics of the firms or establishments (*e.g.*, industry and size).

Identifying establishment closures. I identify establishment closures that occurred between 2009 and 2017. An establishment is treated if it closed or downsized by 90% or more. Additionally, I account for false closures, such as relocations that involve a change in establishment identifier but do not constitute true closures. Using worker flows, I classify a closure as false if more than 70% of the workforce moves to a single new workplace (following [Cestone et al. \[2023\]](#)). I only consider establishments with more than 5 employees 3 years before closure.¹⁴ I drop small establishments because the health of one worker could have an impact on the economic performance of the firm. This would be a threat to the identification strategy because job loss needs to be exogenous to workers' health. I identify 218,559 treated establishments between 2009 and 2017 (Table 1.D.1).¹⁵

¹²This synthetic indicator is developed by the French Ministry of Labor (*DARES*) and the Employment Agency (*Pole Emploi*) and considers three dimensions: the recruitment difficulties anticipated by employers; the number of job offers compared to the number of job seekers; the ease with which job seekers find a job. There is very little variation over time so I use the 2018 indicator at the *département* level. See [DARES and Pole Emploi \[2020\]](#) for details.

¹³I use the Local Potential Accessibility (LPA) indicator developed by the Institute for Research and Documentation in Health Economics (*IRDES*) and the French Ministry of Health (*DREES*). This indicator measures the supply of and demand for general practitioners (GP) by considering practitioners' volume of activity on the one hand, and service use rates differentiated by population age structure on the other. This is a local indicator calculated at the municipal level, it also takes into consideration supply and demand factors in neighboring municipalities. There is very little variation over time so I use the 2015 indicator at the municipality level. See [Barlet et al. \[2012\]](#) for details.

¹⁴This restriction eliminates possible cases of plant closures caused by the poor health of workers. [Leroutier and Ollivier \[2023\]](#) shows, for example, that air pollution has an impact on the health of workers which is reflected in firms' performance. The effect remains small, however, and it seems unlikely that this could lead to plant closures. Nevertheless, the reverse causality link has been established. This restriction excludes 84% of closures.

¹⁵So far, I have identified plant closures at the yearly level, but I plan to redo the analysis at the quarterly level to capture even shorter-term effects.

Matching with *Constances* cohort. In the *Constances* cohort, I observe the identifier of the establishment in which an individual works. I match the two datasets at the establishment level. I can then identify individuals who were employed in an establishment that closed.

2.2.3 Definition of treatment and control groups

The analysis sample is constructed by identifying displaced workers and by selecting a group of workers with similar employment histories prior to job loss. Following the literature on job displacement [Couch and Placzek, 2010; Jacobson et al., 1993], I consider workers with a strong labor market attachment who leave a stable job because of an establishment closure. These criteria select individuals who would probably have kept their jobs if the establishment had not closed.

Restrictions. For each year c , the sample is restricted to all workers who satisfy the following restrictions: being between 25 and 60 years old, being in employment in year $c - 1$, with primary job at c at an establishment which had at least 5 employees in year $c - 3$, and with at least 3 years of tenure.^{16,17}

Treated individuals. An individual is treated (displaced) during year c if the individual leaves the establishment between years c and $c + 1$ and the establishment closes between c and $c + 2$. I thus also include in the treated group the workers who left the workplace one year or two years before the actual closure to account for anticipation effects. If a worker experiences multiple establishment closures over the period, I keep the first occurrence. I identify 20,449 treated individuals (Table 1.D.1).¹⁸

Anticipation of closure. I include in the treated group individuals who left the establishment the year of closure or in the two years preceding closure (see Figures 1.D.1 and 1.D.2 for the composition of the treated group by year of treatment and the timing of employees' exit before a closure).¹⁹ Considering plant closure as an unexpected

¹⁶Individuals are treated when they are between 25 and 60 but I consider all the periods when workers are between 18 and 65 years old. I remove the periods after 65 years old because the effect on health might be caused by age or retirement instead of plant closure.

¹⁷The restriction on the employment status is based on the annual employment status defined in Appendix 1.B. It means that the worker spent most of year $c - 1$ in employment.

¹⁸These restrictions strengthen the internal validity of this paper however it only studies a specific case of job loss.

¹⁹Figure 1.D.1 shows the timing of exit from a firm which will close eventually: 46.9% of employees leave the establishment the year of closure, 21.6% leave one year before closure, 10.0% leave two years before closure and the rest leave before. Figure 1.D.2

event is not realistic, as workers might leave the establishment before the closure occurs. Workers may adapt their job search behavior upon the announcement preceding the establishment closure and leave the establishment because they found a new job. Firms can start separating from workers with mass layoffs before the final closure. In that sense, considering only the workers who were still there during the year of closure could create a selection bias. However, the direction is hard to tell because in the first case, the early displaced workers would be the more re-employable ones and in the second case, the early displaced workers would be the less productive ones.

Matching procedure to select a control group. Working in a closing establishment is non-random. I follow the literature pioneered by [Jacobson et al. \[1993\]](#) and select a matched control group of never-treated workers.²⁰

For each baseline year c , the pool of potential controls is made of all workers following the above restrictions and additionally, they have to be mainly employed in years $c - 1$, c , $c + 1$. This means that they are allowed to lose their jobs (not due to an establishment closure however) in later years but just not at the same time as treated workers.²¹

Matching procedure. I perform a matching procedure to obtain a comparison group with similar trends as the treated group. The trends of this control group represent the counterfactual trends of what would have happened to the treated workers had they not experienced an establishment closure. For each baseline year c , I match an individual treated in year c to an individual among the pool of potential controls with similar employment history and health conditions using propensity score matching.²² I match on i) exact 1-digit industries, age (three-year bins), sex, region and lagged employment status (years $c - 2$ and $c - 3$) and on ii) the propensity score where the predictors are establishment size (size at year $c - 3$), lagged earnings (year $c - 1$), job tenure, occupation and lagged health variables (number of visits to GP, probability to have a hospital stay, probability to take antidepressants and anxiolytics at year $c - 2$).

shows the composition of the treated individuals: 62% left the establishment the year of closure, 29% left one year before closure and 9% left two years before closure.

²⁰I plan to refine this analysis by changing the control group to workers 'not treated this year' instead of 'never treated.' Thus, the control group will consist of both later-treated and never-treated workers. This adjustment aims to create a more convincing control group and avoid selecting individuals conditioning on the post-treatment period.

²¹There is no consensus in the literature on the restrictions in the post-period for control workers. [Jacobson et al. \[1993\]](#) impose that control workers stay employed during the whole post period. [Schmieder et al. \[2022\]](#) do not impose any restriction.

²²I run two alternative matching algorithms. I conduct multivariate-distance matching on characteristics instead of propensity score matching and I estimate propensity scores within each cell (defined by the exact characteristics). The results are similar to the main matching algorithm.

I perform one-to-one matching without replacement so that each treated worker is matched to a single control worker (non-displaced worker with the closest propensity score) and a control worker cannot be used more than once. If a treated worker is matched with multiple controls, one single control worker is randomly selected. I matched 20,449 displaced workers to the same number of control workers (Table 1.D.1).

Analysis sample. All the years of observations of the displaced and control workers are used in the analysis (resulting in 524,633 individual \times year observations). For a treated worker, the baseline year is equal to the year of treatment (closure of the establishment or early exit). For a control worker, the baseline year is equal to the treatment year of the treated individual to whom he is matched.

2.2.4 Descriptive statistics

Comparison of the treated and the control groups. Table 1.1 presents the average characteristics of the workers in terms of demographics, employment history and health outcomes two years before job loss (and placebo event for the control workers). Prior to displacement, displaced workers exhibit different characteristics than the pool of non-displaced workers (almost all means are statistically different, column 5). Displaced workers are more likely to be male, be less experienced and work in smaller firms for example (columns 1 and 3). The matching procedure has successfully selected never-treated individuals: they are reasonably similar in terms of labor market history and healthcare utilization before job loss (means not statistically different, column 4).²³

The treated group is made of individuals aged 39.4 on average two years before job loss and 43.3% are female. They had been employed at the establishment for 6 years on average. The most represented industries are Wholesaling and Retailing/Transport and Storage/Accommodation and Restaurants with 25.6% of workers and Scientific and technical activities/Administrative and support services with 20.9% of workers (Table 1.E.1). 26.4% of displaced workers were white collars and 27.1% were blue-collar workers. They visited the GP on average 3.4 times in the year and 14.7% were taking antidepressants or anxiolytics two years prior to job loss.

²³One can however notice that displaced workers have been employed more years and have spent fewer years in unemployment than matched control workers.

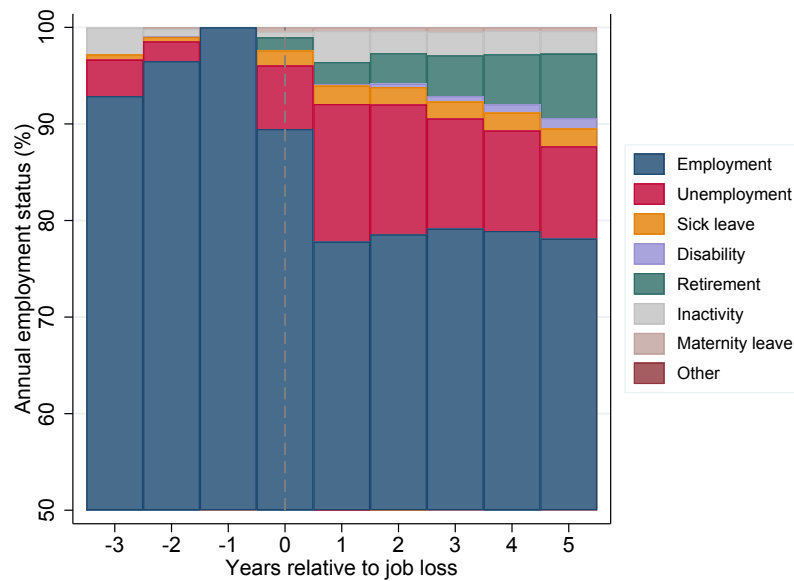
Table 1.1: Summary statistics (year c-2)

	(1) Treated	(2) Matched controls	(3) Random controls	(4) p-value (1)-(2)	(5) p-value (1)-(3)
Panel A: Demographics					
Female (%)	43.28 (49.55)	43.28 (49.55)	54.90 (49.76)	1.000	0.000
Age	39.41 (9.96)	39.39 (9.91)	42.79 (7.30)	0.829	0.000
Panel B: Job characteristics					
Job tenure (in years)	6.02 (5.50)	6.02 (5.37)	3.42 (6.38)	0.932	0.000
Years in employment	8.67 (4.42)	8.82 (4.44)	15.87 (7.73)	0.001	0.000
Annual earnings (2002-euros)	26,790.29 (26,197.06)	27,337.31 (20,552.68)	27,337.60 (23,276.52)	0.019	0.020
Workplace size	291.64 (918.06)	286.57 (965.53)	553.65 (1537.60)	0.588	0.000
Panel C: Health outcomes					
Nb. visits to GP	3.38 (3.71)	3.34 (3.51)	3.77 (3.82)	0.253	0.000
Visits to psychiatrist (%)	3.29 (17.83)	3.11 (17.35)	5.13 (22.07)	0.298	0.000
Antidepressants (%) and anxiolytics	14.75 (35.46)	14.56 (35.27)	17.36 (37.88)	0.586	0.000
Pain killers (%)	56.41 (49.59)	56.43 (49.59)	58.75 (49.23)	0.968	0.000
Hospital stays (%)	11.80 (32.27)	11.57 (31.99)	12.96 (33.58)	0.460	0.000
Observations	20,449	20,449	86,834		

Notes: This table presents summary statistics two years before (placebo) event for several sub-populations. Column 1 is computed for treated workers; Column 2 for matched control workers and Column 3 for the pool of potential control workers (matched and unmatched control workers that satisfy the baseline restrictions). Column 4 reports the p-value from the test for equality of means between column 1 and column 2. Column 5 reports the p-value from the test for equality of means between column 1 and column 3.

Relevance of the shock. I identify exogenous job losses using establishment closures. This exogenous shock is indeed correlated to job loss and employment status. By construction, one year before the closure of the establishment, all treated individuals are employed in the sense of the "annual employment status". One year after closure, 76.7% of treated individuals are employed and 14.6% are unemployed (Figure 1.1). Hence, establishment closures indeed cause job loss with some individuals remaining unemployed for some years and others managing to find another job shortly.

Figure 1.1: Employment status and time to event



Notes: This Figure plots the employment composition of the treated group in each year relative to establishment closure (year 0). The procedure to define the main employment status is described in Appendix 1.B.

3 Empirical strategy

3.1 Main specification

Identification strategy. I use establishment closures as a source of exogenous variation and I implement a combined matching / difference-in-differences approach as identification strategy. I focus on plant closures instead of mass layoffs because they result in all workers losing their jobs regardless of their health condition. I then compare the health evolution of the displaced workers to a matched control group of workers from surviving establishments.

Specification. I exploit variation in exposure to establishment closure across workers. My strategy relies on comparing displaced workers to matched non-displaced workers, before and after the event. The baseline specification is a difference-in-differences event study design with a window of five leads and eight lags. The analysis covers the period from 2007 to 2019, with staggered events occurring between 2009 and 2017. Following recent papers on job loss [Bertheau et al., 2022; Gerard and Naritomi, 2021; Schmieder et al., 2022], I estimate the following equation on the sample of treated and

(matched) control workers.

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=\underline{T} \\ k \neq -3}}^{\bar{T}} \gamma_k D_{it}^k + \sum_{\substack{k=\underline{T} \\ k \neq -3}}^{\bar{T}} \delta_k D_{it}^k \cdot D_i + \epsilon_{it} \quad (1.1)$$

where Y_{it} is the labor/health outcome of individual i in year t . D_i is a dummy equal to 1 if individual i is in the treated group. D_{it}^k are distance-to-event dummies (equal to 1 if year t is k years apart from the event year), formally $D_{it}^k = \mathbb{1}\{t = E_i + k\} \forall t \in (\underline{T}, \bar{T})$, $D_{it}^{\bar{T}} = \mathbb{1}\{t \geq E_i + \bar{T}\}$ and $D_{it}^{\underline{T}} = \mathbb{1}\{t \leq E_i + \underline{T}\}$ where $\bar{T} = 9$, $\underline{T} = -6$ and E_i is the baseline year (year of event for treated individuals and placebo year for control individuals) of individual i . α_i are individual fixed effects, controlling for individual time-invariant characteristics. λ_t are year fixed effects, controlling for time-varying shocks. ϵ_{it} is an error term. Standard errors are clustered at the individual level. The coefficients δ_{-3} and γ_{-3} are normalized to 0.

The parameters of interest are the δ_k 's that correspond to the dynamic impact of job loss k years after ($k \geq 0$) an individual experiences an establishment closure. They measure the difference between displaced workers and their matched controls after taking out the initial difference in year -3.²⁴ They are interpreted as average treatment effects on the treated (ATT).²⁵ The effect is causal under the assumption that, in the absence of event, the (health) outcomes would have evolved similarly for treated and control individuals. This assumption cannot be tested but one can check that the parallel trend assumption holds on the pre-treatment period, *i.e.*, I check that for all $k < 0$ the δ_k 's are not statistically significant. It is important to also control for time-to-event dummies (coefficients γ_k) especially when considering labor market outcomes. The selection criteria (being employed in year -1, with at least 3 years of tenure) imply that treated workers are on an upward work profile before treatment. As workers in the control group are selected if they have a similar employment history to the treated workers, they only appear on this upward profile before the baseline year and this cannot be captured by λ_t alone (see Appendix 1.F.1 for details).

²⁴The reference year in the analysis is $k = -3$ so δ_k measures the difference in outcomes variables between displaced workers and control workers in year k relative to the difference in the reference year -3.

²⁵These estimates could be considered as intention-to-treat (ITT) since only 19% of workers experienced at least one quarter in unemployment following establishment closure (Figure 1.1). Thus the estimated effects relative to the population actually affected by the unemployment transition would be higher.

Aggregated treatment effects. For the sake of readability, I also run static two-way fixed effect models, where I pool together pre and post-event periods. The difference with the main specification is that time-to-event dummies are aggregated into a post-event dummy.

$$Y_{it} = \alpha_i + \lambda_t + \gamma \mathbb{1}\{t \geq E_i\} + \delta \mathbb{1}\{t \geq E_i\} \cdot D_i + \epsilon_{it} \quad (1.2)$$

where $\mathbb{1}\{t \geq E_i\}$ is a dummy equal to 1 if period t is after the event. The parameter of interest is δ .

Heterogeneity analysis. Treatment effects may be heterogeneous and some populations may be more affected by job loss. I run heterogeneity analyses on several dimensions: demographics (sex and age), job characteristics (occupation, earnings, and size of the establishment), local labor market characteristics and local healthcare supply. Let's consider a dimension of heterogeneity X with J groups (*e.g.*, age). Each individual is assigned to a group j invariant with time.²⁶ For the sake of readability, I run a static two-way fixed effect model. For a characteristic X with J modalities, I run the following separate regressions for each sub-population $j \in 1, \dots, J$.

$$Y_{it}^j = \alpha_i^j + \lambda_t^j + \gamma^j \mathbb{1}\{t \geq E_i\} + \delta^j \mathbb{1}\{t \geq E_i\} \cdot D_i + \epsilon_{it}^j \quad (1.3)$$

where Y_{it}^j is the health outcome at year t of individual i being in group j , $\mathbb{1}\{t \geq E_i\}$ is a dummy equal to 1 if period t is after the event. I plot the parameter of interest δ^j for each sub-population $j \in 1, \dots, J$.

3.2 Alternative specifications and robustness checks

The recent literature on two-way fixed effect (TWFE) models has shown that this specification may lead to biased estimators in the presence of heterogeneous treatment effects over time and between groups. The standard TWFE estimator estimates weighted sums of average treatment effects with potentially negative weights, which can result in a negative estimate whereas all treatment effects are positive and vice versa [[de Chaisemartin and D'Haultfoeuille, 2020a](#); [Goodman-Bacon, 2021](#)]. Regarding job

²⁶The groups are generally defined according to their characteristics the year preceding closure.

loss, treatment effects are likely to be heterogeneous both over time and between individuals. Additionally, even if there is no treatment effect heterogeneity across cohorts, the TWFE coefficients can be contaminated by treatment effects from other periods [Sun and Abraham, 2021]. Many papers have proposed new alternative estimators to address the issues mentioned above [Borusyak et al., 2022; Callaway and Sant’Anna, 2021; de Chaisemartin and D’Haultfoeuille, 2020b; Goodman-Bacon, 2021; Sun and Abraham, 2021]. They offer alternatives to the standard TWFE estimators which are unbiased even when the effect of the event varies over groups and periods. However, in the current design, these complications might be avoided thanks to the presence of never-treated workers and their selection (matched treatment-control groups for each displacement year).²⁷ I perform the robustness checks described below on the specification and on the definition of the control group and I include the appropriate alternative estimators in the main graphs.

Specification. As explained above, distance-to-event dummies are necessary to account for an upward work profile before baseline year.²⁸ However, this comment is especially true for labor market outcomes but less so for health outcomes (because treated and matched individuals are not selected based on their health outcomes). I run robustness checks where I remove distance-to-event dummies for control individuals and I also implement the alternative estimators designed by Sun and Abraham [2021] and Callaway and Sant’Anna [2021] which use never-treated individuals. The estimation and results are presented in Appendix 1.F.2. Results on health outcomes are consistent with the main specification.

Control group. In the main specification, I define a control group of never-treated individuals selected on observables. As robustness checks, I only consider ever-treated individuals so the control group are the not-yet-treated individuals and I also implement the alternative estimator of de Chaisemartin and D’Haultfoeuille [2020b] which relies on the comparison of already-treated individuals to not-yet-treated individuals. The estimation and results are presented in Appendix 1.F.2. Results on health outcomes are

²⁷The presence of a large group of untreated units puts less weight on the “problematic” 2x2 comparisons that use already treated units as controls.

²⁸See also detailed discussion in Schmieder et al. [2022] and explanations in Appendix 1.F.1.

consistent with the main specification.

Anticipation effects and gradual downsizing. Plant closures may be anticipated because they are usually preceded by mass layoffs. Announcements of those can have an impact on workers' job search behavior [Cederlöf et al., 2021]. In the main specification, I also include in the treatment group those who left the firm one year or two years before the final closure. I run a robustness check where the treatment group is only made of workers who left the firm in the year of closure.

4 Results

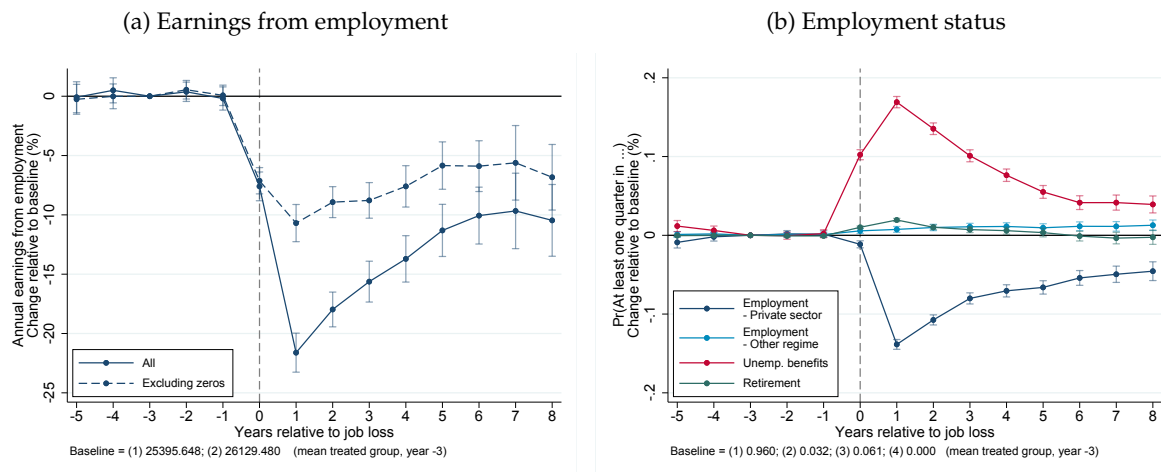
4.1 Effect of job loss on Economic Outcomes

I begin by providing evidence on the long-term impact of job displacement on a range of economic outcomes (annual earnings, employment status and probability of receiving unemployment benefits) up to eight years after job loss.

Displaced workers experience a sharp drop in annual earnings and they slowly recover in subsequent years. Figure 1.2a compares the evolution of earnings for treated and (matched) control groups relative to year $c - 3$. Earnings are 5,000 euros lower in the year following job loss for the displaced workers (20% reduction relative to year $c - 3$ mean) and they remain 3,000 euros lower five years after displacement (12% reduction relative to year $c - 3$ mean). Even those who find a new job receive lower incomes (on average 2,000 euros lower four years after displacement). This large drop in earnings is comparable to those found in the US [Jacobson et al., 1993; Sullivan and von Wachter, 2009] and in Europe (Schmieder et al. [2022] in Germany, Roulet [2017] in Denmark, Bertheau et al. [2022] for France). Figure 1.2b shows that a huge part of the earnings losses is explained by the extensive margin (change in employment status). The probability of being employed at least one quarter drops by 14pp (15% reduction relative to year $c - 3$ mean) in the year following job loss compared to the control group but recovers faster than earnings. More than five years after job loss, displaced workers are 7pp less likely to be employed at least one quarter than control workers. The probability of receiving unemployment benefits increases by 17pp for

the displaced workers in the year following job loss compared to the control workers and is still 6pp higher five years after job loss.

Figure 1.2: Effect of job loss on economic outcomes



Notes: These figures plot the δ_k 's coefficients obtained from the staggered difference-in-difference regressions (equation 1.1). δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -3) between the treated group and the control group. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Outcomes: (a) Annual earnings from employment; (b) Probability to be employed (private sector or other), to receive unemployment benefits, and to be retired. The standard errors are clustered at the individual level.

Heterogeneity analysis. Workers are not all affected in the same way by the closure of the firm they worked in. Men, older workers, workers in small businesses and low-skilled workers are suffering the biggest drops in earnings. Additionally, low-paid workers are more likely to receive unemployment benefits (Figure 1.G.1).

4.2 Effect of job loss on Mental Health Outcomes

I now investigate whether job displacement has an effect on health and I focus on mental health outcomes. I consider a range of mental health outcomes and I study the effect up to eight years after displacement. I look at several healthcare categories: visits to doctors, hospital stays and prescribed drugs.

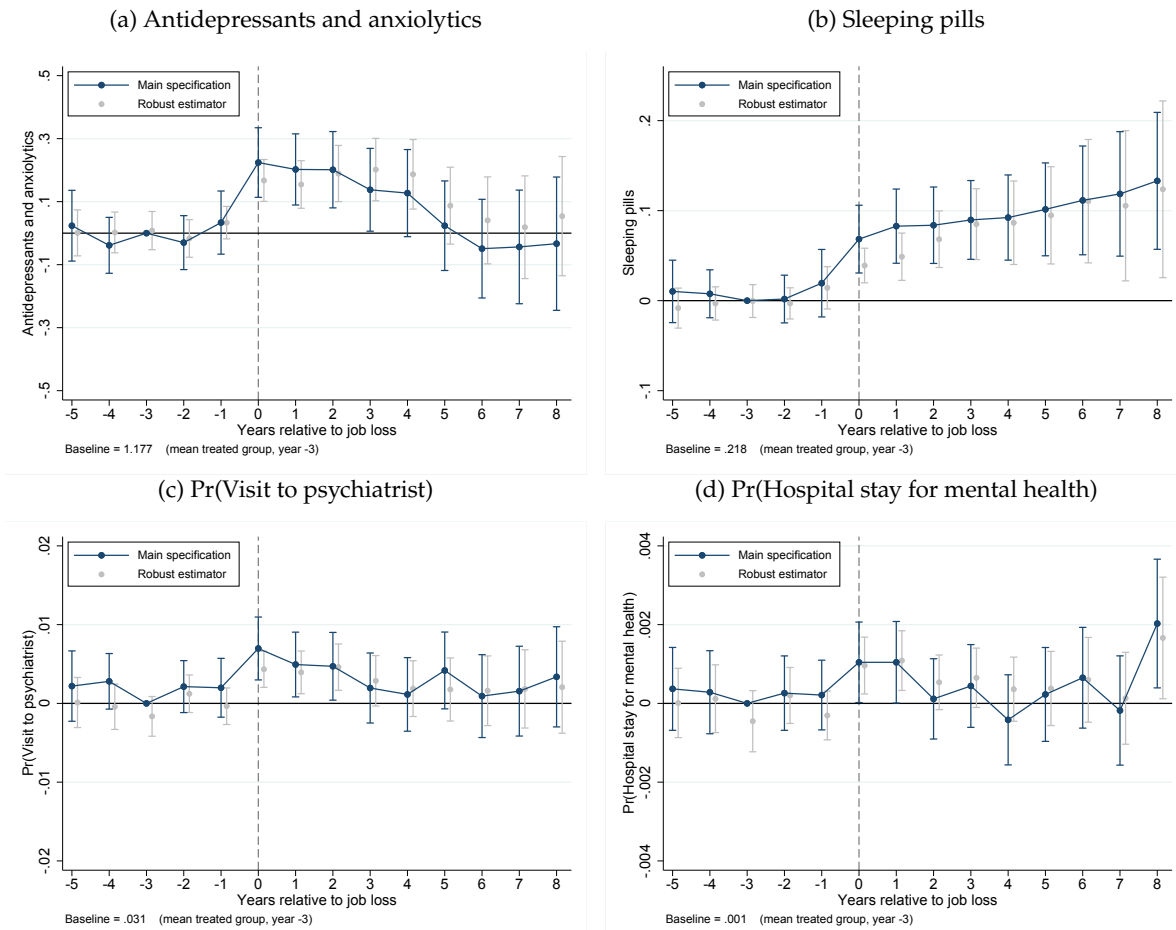
Prescribed drugs. I look at prescribed drugs related to mental health: antidepressants, anxiolytics and sleeping pills. The use of antidepressants and anxiolytics is 0.2 boxes higher for displaced workers (19% increase relative to $c - 3$ mean) in the four years following job loss (Figure 1.3a) and the difference becomes insignificant afterward. The use of sleeping pills is 0.5 boxes higher for displaced workers (32% increase relative to $c - 3$ mean) in the year of displacement and the effect increases even up to eight

years after displacement (Figure 1.3b). This may be explained by the addictive effect of sleeping pills. Even if the use of sleeping pills starts increasing at job loss, the cumulative effect is even greater several years later. For both types of prescribed drugs, one can notice a small anticipation effect the year before displacement which can be due to the stress generated by the announcement of closure or the financial distress of the firm.

Visits to doctors and Hospital stays. I then look at visits to psychiatrists and hospital stays. The likelihood of visiting a psychiatrist increases by 0.5pp (20% increase relative to $c - 3$ mean) in the three years following job loss (Figure 1.3c). The effect stays positive but not significant up to eight years after job loss. There is a positive and significant effect on hospital stays for mental health (Table 1.H.2) in the two years after job loss: the probability of having a hospital stay with this diagnosis issues double in the year of displacement. The effects on hospital admissions and visits to doctors are less persistent than for prescribed drugs.

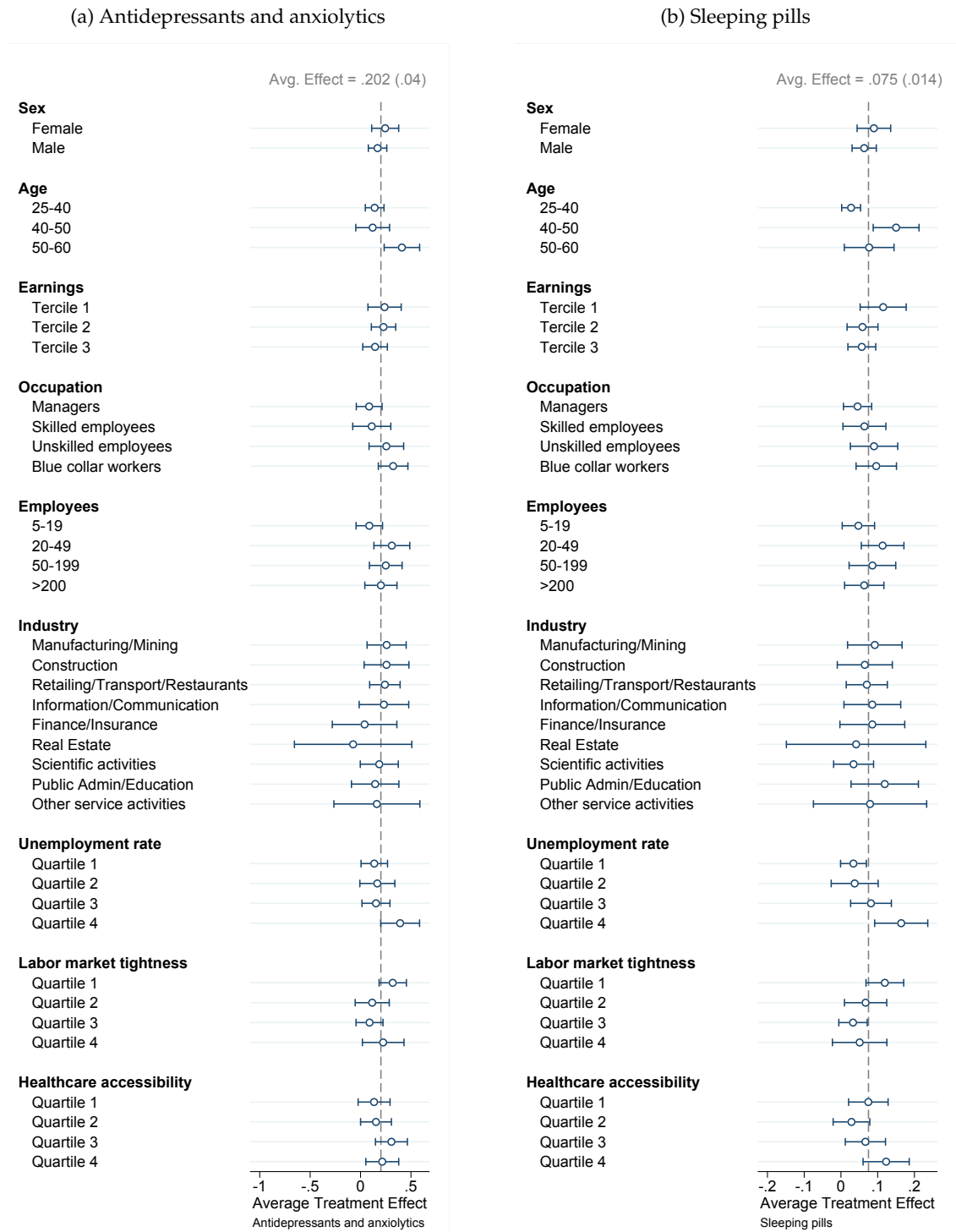
Heterogeneity analysis. Older workers (above 50) and workers living in places with a high unemployment rate experience the highest increase in the use of antidepressants and anxiolytics (Figure 1.4a). The effect is significant for unskilled employees and blue-collar workers but not for managers and skilled employees. One can notice that the effect is only significant for two industries: retailing/transport/restaurants and scientific activities.

Figure 1.3: Effect of job loss on mental health



Notes: These figures plot the δ_k 's coefficients obtained from the staggered difference-in-difference regressions (equation 1.1) and the alternative estimator of Callaway and Sant'Anna [2021]. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -3) between the treated group and the control group. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Outcomes: (a) Number of boxes of antidepressants and anxiolytics; (b) Number of boxes of sleeping pills; (c) Probability to visit a psychiatrist; (d) Probability to have a hospital stay for mental health. The standard errors are clustered at the individual level.

Figure 1.4: Heterogeneity (mental health-related drugs)



Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 1.3). The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. The standard errors are clustered at the individual level.

4.3 Effect of job loss on other Health Outcomes

Job loss can also have consequences on physical well-being and affect behaviors toward health coverage. To complement the main results, I also look at other outcomes

related to physical health (visits to doctors, hospital stays and prescribed drugs) and health insurance (recognition of long-term conditions and disability benefits). All graphs and tables are presented in Appendix 1.H.

Visits to doctors. I first look at visits to GPs and specialists. Job loss results in an increase in the number of visits to GP by 0.2 (6% increase relative to year $c - 3$ mean) in the year of closure only (Figure 1.H.1a). Visits to gynecologists, rheumatologists, anesthetists, and neurologists significantly increase and visits to ophthalmologists significantly decrease in the year of job loss and the year after (Table 1.H.1).

Hospital stays. I then look at hospital stays with some specific diagnosis. Job loss significantly increases the likelihood of having a hospital stay by 13% in the year of displacement (Figure 1.H.1b). The effect is driven by hospital stays for cancer diagnoses and maternity and is not significant for diagnoses related to circulatory problems, alcohol-related issues nor for the use of the emergency department (Table 1.H.2). The findings regarding cancer hint at a potentially positive aspect of job loss: individuals may have more time to take care of their health. It is unlikely that job loss would lead to the development of cancer in such a short time. People could have postponed their health check-ups, such as cancer screening, while they were working and taken them once they had more time available.

Prescribed drugs. I look at the consumption of painkillers, opioids and anti-inflammatories. [Case and Deaton \[2015\]](#) attributed the rise in mortality rates among middle-aged white men and women in the United States in the 2000s to increased consumption of painkillers, particularly opioids in the United States ("deaths of despair"). In France, the scale of this phenomenon is much smaller. Yet, it appears that job loss does increase long-term opioid use. Job loss increases the use of painkillers by 20% and of opioids by 24% five years after job loss (Figures 1.H.1c and 1.H.1d). The effect is not significant for anti-inflammatory drugs.

Long-term conditions. Displaced workers are also more likely (probability increases by around 10%) to be recognized with long-term conditions up to five years after

job loss (Table 1.H.3a).^{29,30} The conditions with significant increase are stroke, artheriopathies, immune deficiency, Parkinson's disease, psychiatric disorders and malignant tumors (Table 1.H.3).

Heterogeneity analysis. The effect on the number of visits to GP and the use of painkillers is significant for workers above 50 (Figure 1.H.2). One can also notice that the use of painkillers increases for unskilled employees and blue-collar workers but decreases for managers, suggesting that the difference in tasks may play a role in the consequences of job loss.

5 Channels

Section 4.2 provides evidence of the detrimental effect of job loss on mental health and I now examine the potential causes.

The first channel is the income loss which can generate stress related to economic insecurity.³¹

New job opportunities. A first mechanism that could explain the deterioration in displaced workers' mental state is that they are worried about the loss of income and the difficulty of finding a new job. If this prediction holds then new job opportunities should mitigate the deterioration in mental health. To capture these reemployment perspectives I compute an individual risk on the labor market. I do not use the realized employment status because it is endogenous. Instead, I measure individual risk using the probability of finding a new job in the following year, which also captures the expected decline in earnings. To do so, I estimate re-employment probabilities using leave-one-out predictions (based on demographic characteristics, employment history and health conditions; see Appendix 1.I for details). High-risk workers experience a significant increase in the use of mental health drugs, while the effect is not significant

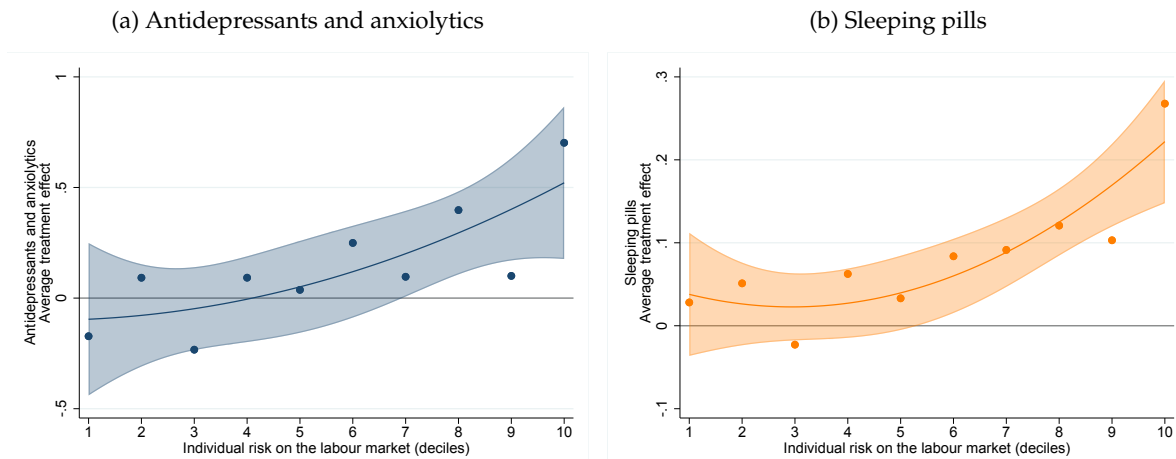
²⁹In France, being recognized with specific long-term conditions makes you eligible to specific healthcare reimbursement with full coverage

³⁰I also look at the impact on disability benefits. Job loss has a positive and significant impact on the probability of receiving disability benefits. One year after job loss, the probability of receiving disability benefits is 4pp higher for displaced workers and this effect is increasing over time (Figure 1.H.3a). Older workers (aged 50-60) and low-earning workers experience the highest increase in the probability of receiving disability benefits (Figure 1.H.4). However, we do not know whether this is due to a change in eligibility or in claims only. But this likely reflects a change in incentives: displaced workers may have replaced income from employment with income from disability benefits.

³¹Healthcare renunciation could also be a mechanism but is less likely because the French universal health insurance reduces the relative cost of healthcare.

for middle to low-risk workers (Figure 1.5). The effect is concentrated among workers in the three highest risk deciles for the use of antidepressants and anxiolytics and concerns workers above the median risk for sleeping pills. Figures with dynamic treatment effects are displayed in Appendix 1.I. Low-risk workers are not affected by job displacement, even during the years following displacement (Figure 1.I.1b).

Figure 1.5: Individual risk on the labor market and mental health



Notes: These figures plot the δ coefficient from equation (1.2). Separate regressions are estimated by deciles of individual risk. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. The standard errors are clustered at the individual level.

Transition into vs. duration of unemployment. The timing of job displacement can be decomposed into i) the transition into unemployment and ii) the duration of unemployment. The observed decline in mental health may indeed stem from the immediate shift in status resulting from the transition into unemployment or the changes in routine and the depreciation in human capital associated with the time spent not working. The jump in the consumption of antidepressants and anxiolytics occurs directly in the year of displacement in Figure 1.3a and the magnitude of the coefficients stays stable over the two following years. This suggests that the effect might be driven by the transition into unemployment rather than the duration of unemployment. The interpretation might be different for sleeping pills as the size of the effect increases over time.

A second channel is the psychological consequences of the announcement and of changing routine. The mental health outcomes studied above were defined based on administrative data (*e.g.*, use of antidepressants and anxiolytics, visit to psychiatrists) and

they capture severe episodes of depression. However, it can also be expected that job loss will lead to more common episodes of emotional distress that have not resulted in a visit to a psychiatrist or the use of medication. As a next step, I plan to leverage self-reported outcomes in the *Constances* survey data to capture changes in self-assessed health, current state of mind, lifestyle and risky behaviors.³²

6 Conclusion

This paper investigates the causal and dynamic effect of job loss on a comprehensive range of health outcomes in France. Reverse causality and selection bias are challenges here as poor health could be the reason for job loss. To overcome this problem, I follow the strand of literature that identifies exogenous job losses with establishment closure or downsizing. Leveraging administrative data from a French epidemiological cohort, I identify 20,000 exogenous job losses of private sector employees with strong labor market attachment. In a staggered event study, I compare displaced workers to control workers with similar employment histories. While public policies have been designed to mitigate the expected income decline following job loss, little is known about how this income reduction and associated lifestyle changes can affect health. This analysis is conducted in France, where individuals are well insured in terms of employment (with limited income loss risk) and healthcare (relatively low access costs). Yet, findings point towards negative effects on various aspects of health following job loss. Job loss appears to have adverse effects on mental health. Specifically, I find increases in the use of antidepressants, anxiolytics, and sleeping pills, as well as a higher likelihood of psychiatric consultations in the short term. Hospital admissions also exhibit a significant effect. Additionally, job loss leads to a higher probability of receiving disability benefits. Furthermore, treatment effects are heterogeneous with certain vulnerable subpopulations, such as older workers, low-earning workers and workers living in places with high unemployment rates, experiencing more pronounced impacts from job loss. The perspective of re-employment also plays an important role in mitigating

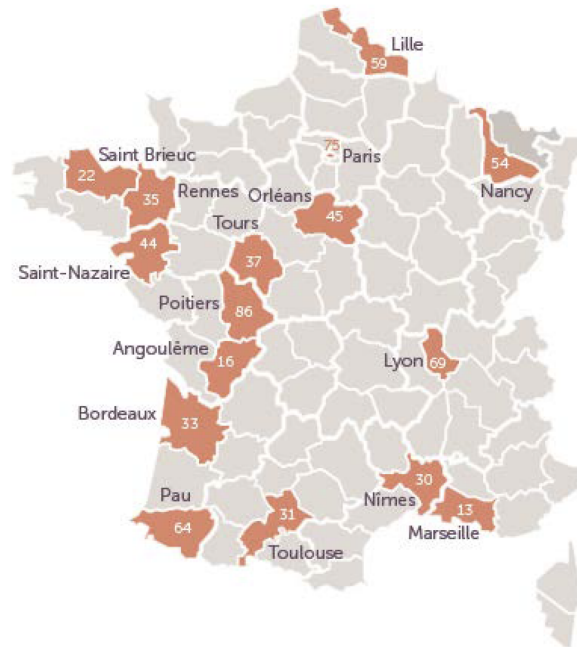
³²Note that the sample size is much smaller due to the survey features: *active participants* should be treated between 2013 and 2017 because the questionnaires start in 2012, and the panel is not balanced because of the potential gaps between follow-up questionnaires.

the consequences on mental health. These results indicate that income replacement for individuals is indeed necessary following job loss, but that they could also benefit from psychological support following this event and in their future jobs.

Appendices to Chapter 1

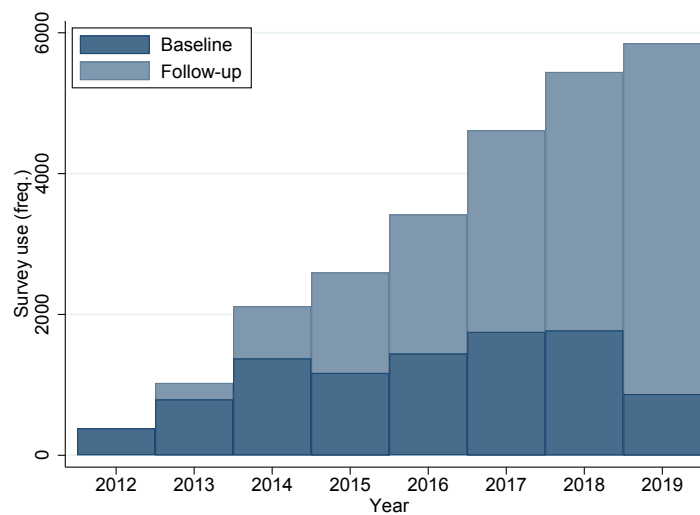
1.A Constances cohort

Figure 1.A.1: Map of targeted zones by Constances



Notes: This map plots the 16 *départements* (with partner Health Screening Centers) where the Constances population was drawn.

Figure 1.A.2: Participation to survey



Notes: This figure plots the number of respondents to the questionnaires per year (distinguishing between inclusion and follow-up questionnaires). The graph is based on all active participants..

1.B Definition of labor market outcomes

Administrative data on labor market outcomes come from the National Retirement Insurance Fund and provide the number of validated quarters that contribute to an

individual's pension.

Quarters validation.

Table 1.B.1: Requirements to validate quarters in each status

Status	Requirements to validate 1 quarter ³³
Employment	Earnings $\geq 200 \times$ gross hourly minimum wage (before 2014) Earnings $\geq 150 \times$ gross hourly minimum wage (after 2014)
Unemployment	50 days of unemployment benefits
Sick leave and work accident	60 days of compensation
Disability	90 days of compensation

Employment status. I define six statuses that correspond to the main employment situation for each year: "Employment", "Unemployment", "Inactivity", "Retirement", "Sick leave and work accident" and "Maternity leave". Following [Rabaté and Rochut \[2017\]](#), the computation is based on

- the number of quarters validated in each state for a given year (hierarchical order in case of a tie: employment, unemployment, sickness, maternity leave, disability, retirement and inactivity)
- with additional correction for time spent in each state when information is available (e.g. retirement date; unemployment, sickness and disability periods correspond to a minimum number of days spent in the state)

1.C Definition of health outcomes

Administrative data on the healthcare consumption of individuals come from the Nation Health Insurance Fund.

³³Requirements to validate quarters are presented here: <https://www.legifrance.gouv.fr/codes/id/LEGISCTA000006173397/>

- Unemployment: https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_chomage_periode_validable_ex
- Sick leave and work accident: https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_maladie_maladie_ex; https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_accident_travail_accident_travail_ex
- Disability: https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_invalidite_invalidite_ex

Table 1.C.1: Definition of health outcomes

Outcome	Definition
Consultations	<i>Visits to</i>
GP	"General practitioners", "Specialist in general medicine with diploma", "Specialist in general medicine recognized by the Order"
Mental health specialist	"Psychiatrists" and "Neuropsychiatrists"
Hospital stays	<i>CIM10 classification</i>
All (except pregnancy)	All except O00-O99
Cancer	C00-D48
Circulatory problems	I00-I99
Alcohol-related issues	F10, I85, K70, K860, T500, T510
Mental health	F00-F99
Drugs	<i>ATC7 classification</i>
Antidepressants and anxiolytics	N06A and N05B
Sleeping pills	N05C
Pain killers	N02
Opioids	N02A
Anti-inflammatories	M01

1.D Definition of treatment

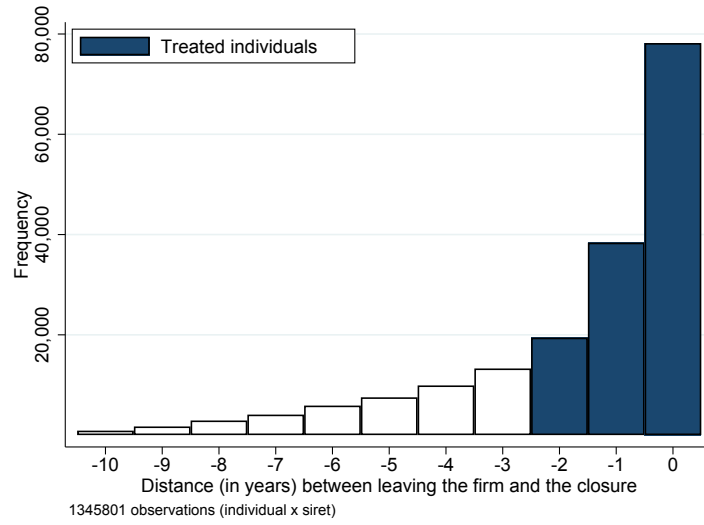
1.D.1 Details on the actual definition of treatment

Table 1.D.1: Explanation of the sample size of the treated group

Establishments (DADS data)	
Nb. closing establishments <i>Closure or ≥ 90% downsizing between 2008 and 2017</i>	1,330,699
Nb. treated establishments <i>5+ employees 3 years before closure</i>	218,559
Individuals (Constances data)	
Nb. individuals at treated establishments <i>Primary job at a treated establishment the year of closure or in the two years preceding closure</i>	69,347
Nb. treated individuals <i>Aged 25-60 Tenure of 3+ years Mainly employed in c-1</i>	21,339
Nb. matched treated individuals	20,449

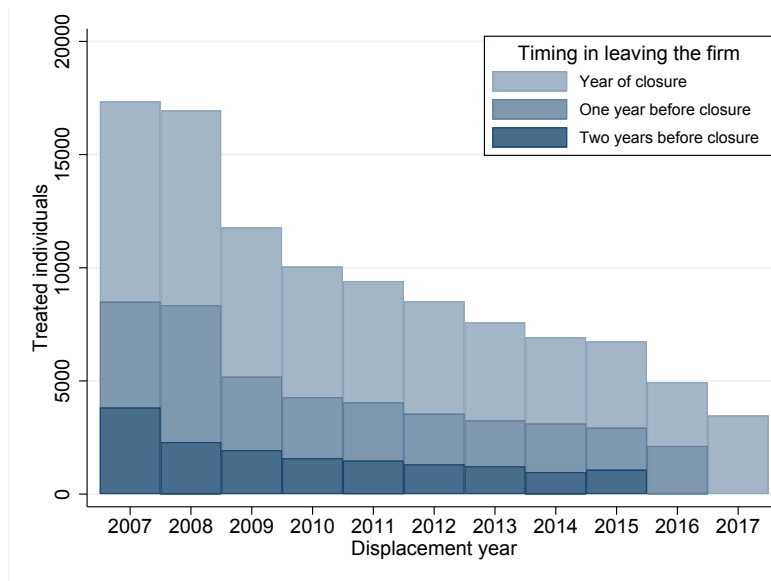
Notes: This table explains the size of the analysis sample. The first panel on establishments shows the number of closing establishments following the restrictions. The second panel shows the number of displaced workers following the restrictions.

Figure 1.D.1: Timing of employees' exit before closure



Notes: This figure plots the distance (in years) between the closure of an establishment (time 0) and the departure of an employee. The graph is based on all workers who worked in their life in an establishment that closed eventually.

Figure 1.D.2: Composition of the treated group



Notes: This figure plots the composition of the three types of treated individuals according to the year of treatment. The graph is based on treated workers.

1.E Additional summary statistics

Table 1.E.1: Summary statistics on industry and socio-professional category (year -2)

	Displaced workers
Panel A: Industry	
Wholesaling and Retailing/Transport and Storage/ Accommodation and Restaurants	25.6%
Scientific and technical activities/Admin. and support services	20.9%
Manufacturing/Mining/Other	15.6%
Public Admin./Education/Social work	13.2%
Construction	7.1%
Information/Communication	6.8%
Finance/Insurance	5.0%
Other service activities	4.3%
Real Estate	1.0%
Missing	0.5%
Panel B: Occupation	
Managers	26.4%
Skilled employees	17.2%
Unskilled employees	26.3%
Blue collar workers	27.1%
Missing	3.0%

Notes: This table shows summary statistics for the treated group in the year before closure. The treated group is made of eligible displaced workers. The first panel is at the industry level. The second panel is at the occupation level.

1.F Main specification and robustness checks

1.F.1 Explanation for the main specification

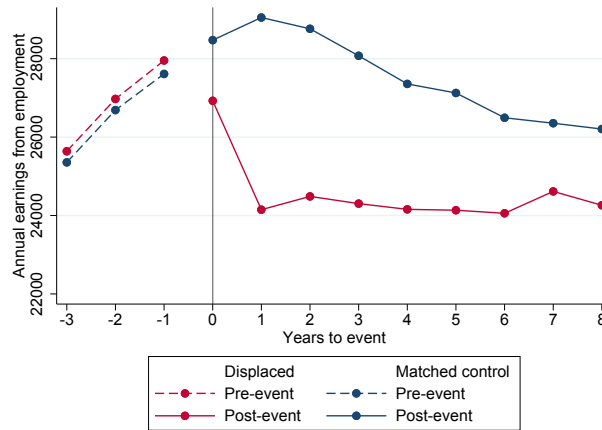
Control group. The selection criteria to define the treated individuals imply that i) treated individuals are on an upward work profile before treatment and ii) work status may deteriorate even without being treated once restrictions are lifted (mean reversion).³⁴ One can for instance notice that earnings decrease for the control group too even if the control workers do not experience a job loss in year 0 (Figure 1.F.1). Therefore, the use of a control group makes it possible to control for the fact that the people treated had experienced a specific career path prior to the event.

Distance-to-event dummies for the control group γ_k . The control group is selected

³⁴The selection criteria is that eligible individuals have at least three years of tenure and were in employment the year before.

with the same criteria as the treated group. Hence control workers also have an upward work profile before baseline. Therefore, I need to control for this specific period in their career.

Figure 1.F.1: Trends in earnings



Notes: This graph plots the average annual earnings from employment in each year relative to baseline for the treated and the control group.

1.F.2 Robustness checks: two-way fixed effect model and only ever-treated

I perform two robustness checks on the specification (Table 1.F.1). First, I use a standard two-way fixed effect model by removing the distance-to-event dummies for the control group. Second, I use ever-treated workers so that the group control is no longer made up of never-treated workers but of not-yet-treated workers.

Table 1.F.1: Robustness checks on the specification

Specification	Main specification	Alternative spe. 1	Alternative spe. 2
Sample	Treated and matched controls	Treated and matched controls	Treated
Estimation			
Distance to event dummies	yes	no	-
× Treated	yes	yes	yes
Time FE	yes	yes	yes
Individual FE	yes	yes	yes

- **Alternative specification 1: Estimation with standard two-way fixed effect model**

I only include distance-to-event dummies for treated individuals (which correspond to a standard TWFE model, without γ_k from the main specification). I estimate the following regressions.

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-10 \\ k \neq -3}}^{11} \delta_k D_{it}^k \cdot D_i + \epsilon_{it} \quad (1.F.1)$$

The coefficients δ_{-3} is normalized to 0.

I also implement the alternative estimators of [Sun and Abraham \[2021\]](#) and [Callaway and Sant'Anna \[2021\]](#) which rely on the comparison of individuals who experienced a job loss to never-treated individuals.

- **Alternative specification 2: Estimation with only ever-treated**

I only consider ever-treated individuals. In this setting, I use later-treated individuals as a control group for early-treated individuals and I exploit the exogeneity in the timing of the event. I estimate the following regressions.

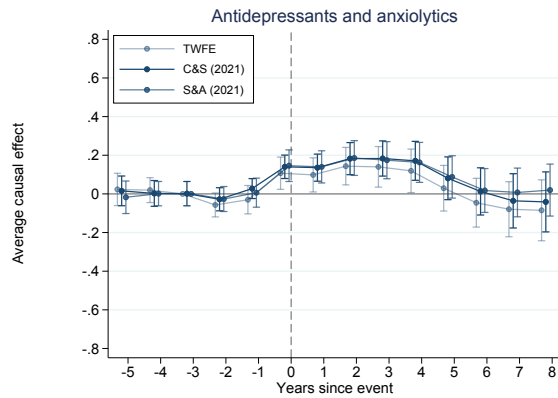
$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-10 \\ k \neq -3}}^8 \delta_k D_{it}^k + \epsilon_{it} \quad (1.F.2)$$

The coefficients δ_{-3} and δ_{-10} are normalized to 0 (with only ever-treated individuals, one needs to omit two periods as explained in [Borusyak and Jaravel \[2017\]](#)).

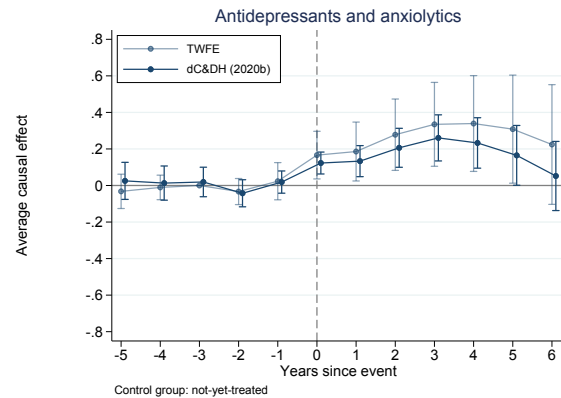
I also implement the alternative estimator of [de Chaisemartin and D'Haultfoeuille \[2020b\]](#) which relies on the comparison of individuals who already experienced a job loss to not-yet-treated individuals.

Figure 1.F.2: Robustness checks using standard TWFE and only ever-treated individuals

(a) Alternative spec. 1 + alternative estimators



(b) Alternative spec. 2 + alternative estimators



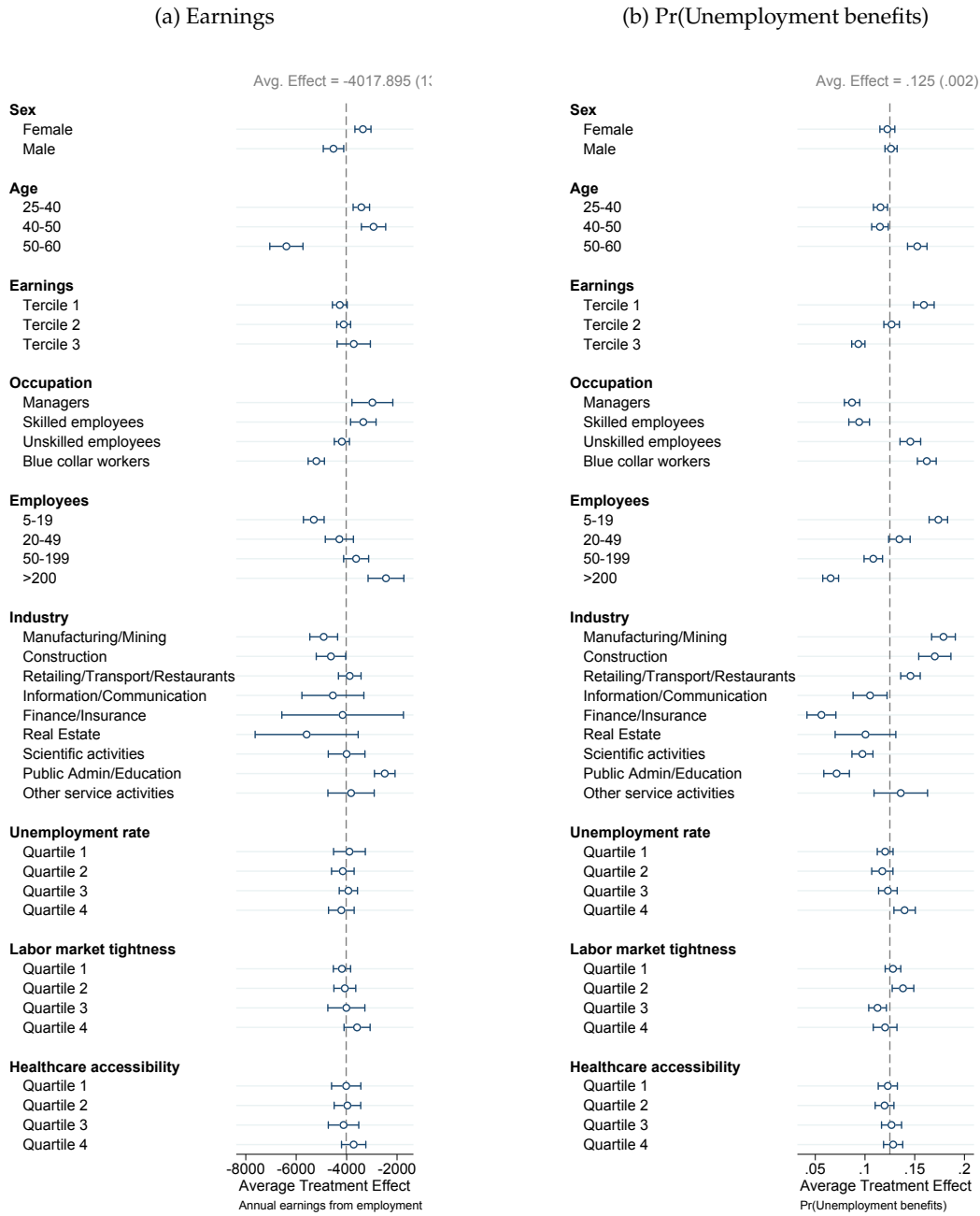
Notes: These figures plot the distance to event coefficients (δ_k) following the event-study regressions (equation 1.F.1). The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Year -3 is taken as the reference year for the TWFE specification. Alternative specification 1 corresponds to the standard TWFE model with a never-treated group; C&S (2021) is the Callaway and Sant'Anna [2021] estimator and S&A (2021) is the Sun and Abraham [2021] estimator. Alternative specification 2 corresponds to the standard TWFE model with only ever-treated individuals; dC&DH (2020b) is the de Chaisemartin and D'Haultfoeuille [2020b] estimator. The standard errors are clustered at the individual level.

Comparison of specifications. Results on health outcomes are very consistent with the main specification when using standard TWFE models and when using not-yet-treated individuals as control group (Figure 1.F.2).³⁵

³⁵Figure 1.F.2 shows only results for the outcome "Antidepressants and sleeping pills" but results are very similar for all health outcomes. Results differ, however, when looking at labor market outcomes because the specification ignores the specific career path of individuals before the event (as explained in section 1.F.1).

1.G Additional figures for labor market outcomes

Figure 1.G.1: Heterogeneity (economic outcomes)

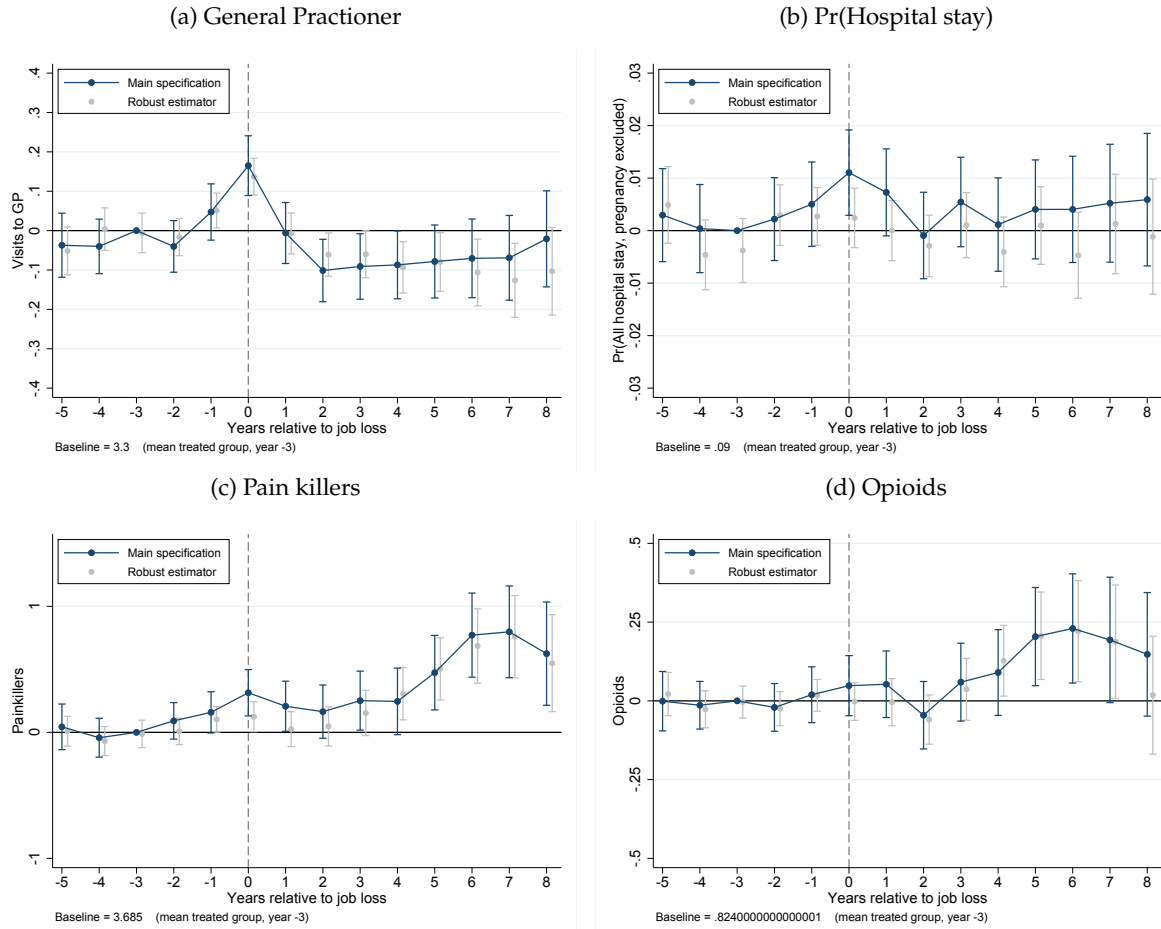


Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 1.3). The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. The standard errors are clustered at the individual level.

1.H Additional figures for health outcomes

1.H.1 Physical health

Figure 1.H.1: Effect of job loss on physical health



Notes: These figures plot the δ_k 's coefficients obtained from the event-study regressions (equation 1.1). δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -3) between the treated group and the control group. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Outcomes: (a) Number of visits to a GP; (b) Probability of hospital admission (excluding pregnancy); (c) Number of boxes of painkillers (d) Number of boxes of opioids. The standard errors are clustered at the individual level.

Table 1.H.1: Disaggregated outcomes for doctors' visits

	0-1 years	2-3 years	4+ years
GP	0.0892*** (0.0260)	-0.0868** (0.0297)	-0.0530 (0.0321)
Psychiatrist	0.0440 (0.0268)	-0.0104 (0.0319)	-0.0171 (0.0376)
Gynecologist	0.0159* (0.00723)	-0.00661 (0.00801)	0.00433 (0.00794)
Cardiologist	0.000333 (0.00221)	-0.00230 (0.00244)	0.00109 (0.00247)
Lung Specialist	-0.00103 (0.00118)	-0.000775 (0.00130)	0.000176 (0.00150)
Dental Facial Surgeon	-0.00128 (0.00137)	-0.000624 (0.00133)	-0.00113 (0.00132)
Oncologist	0.00157 (0.00111)	0.00149 (0.00164)	0.00486 (0.00314)
Rheumatologist	0.0167*** (0.00474)	0.00287 (0.00522)	-0.00463 (0.00530)
Ophthalmologist	-0.00786* (0.00361)	-0.00534 (0.00367)	-0.00154 (0.00366)
Dermatologist	-0.00519 (0.00461)	-0.00697 (0.00485)	-0.00119 (0.00487)
Gastro Urologist	0.00613 (0.0104)	0.0110 (0.0123)	0.00551 (0.00800)
Child Specialist	-0.000563 (0.00297)	-0.000445 (0.00324)	-0.000507 (0.00384)
Anaesthetist	0.00468* (0.00209)	-0.000169 (0.00219)	0.00143 (0.00208)
Endocrinologist	-0.000291 (0.00245)	-0.00397 (0.00285)	-0.00380 (0.00311)
Otorhynolaringologist	-0.00147 (0.00223)	-0.000408 (0.00239)	-0.00340 (0.00252)
Neurologist	0.00404** (0.00148)	0.00195 (0.00176)	0.00227 (0.00179)
General Surgeon	-0.00150 (0.00192)	-0.00130 (0.00199)	0.000503 (0.00202)
Rehabilitation Medicine	0.00262 (0.00201)	0.00113 (0.00245)	-0.000398 (0.00242)
Plastic Surgeon	-0.00131 (0.000934)	-0.000754 (0.00103)	-0.00288** (0.000999)
Internal Medicine	-0.00308 (0.00157)	0.000771 (0.00164)	-0.000209 (0.00240)
Other	-0.000426 (0.000581)	0.000162 (0.000764)	-0.000820 (0.000981)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

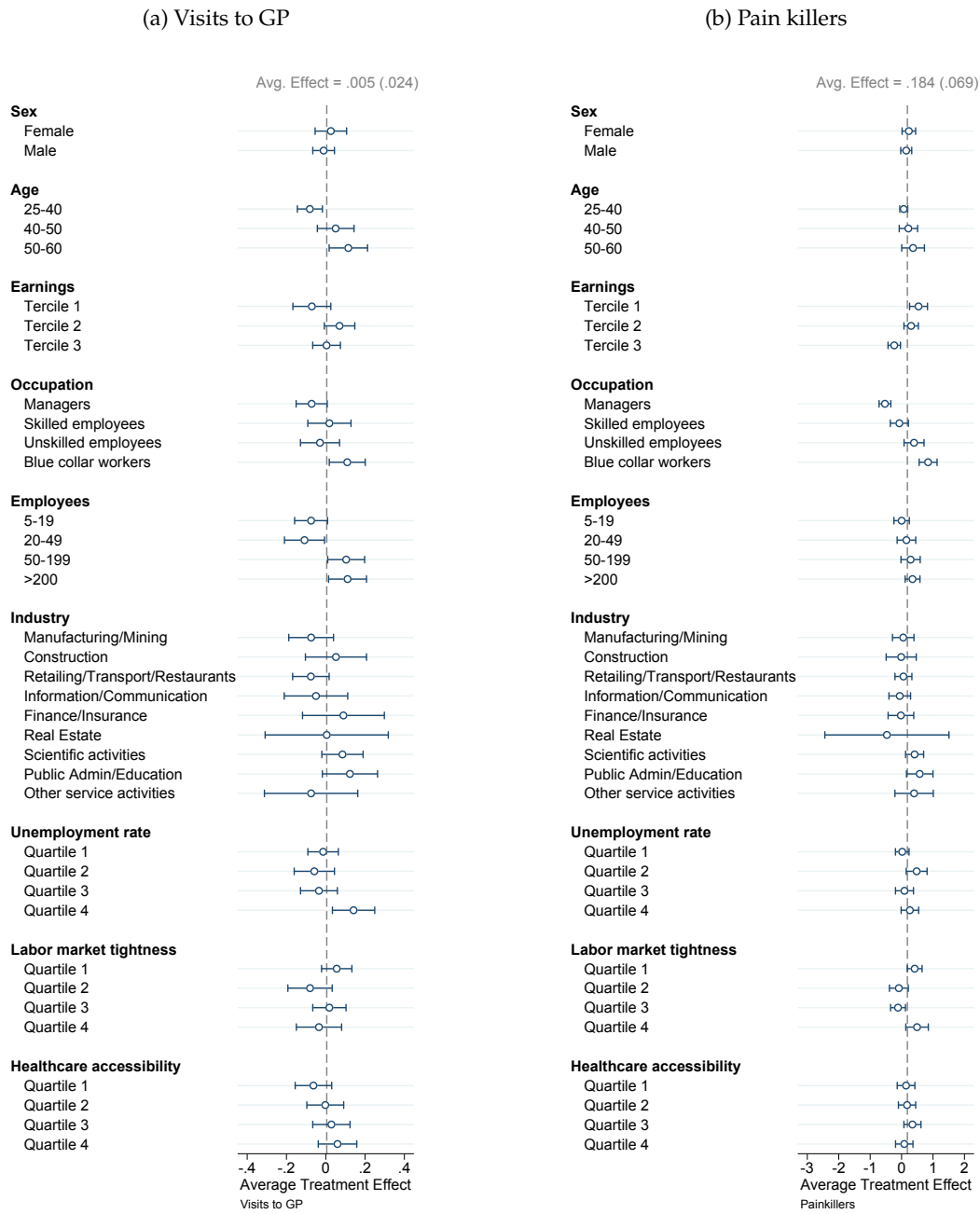
Table 1.H.2: Disaggregated outcomes for hospital admissions

	0-1 years	2-3 years	4+ years
All (except pregnancy)	0.006* (0.002)	-0.001 (0.003)	0.000 (0.002)
Circulatory problems	-0.000 (0.001)	-0.000 (0.001)	0.001 (0.001)
Cancer	0.002* (0.001)	0.001 (0.001)	-0.001 (0.001)
Mental health	0.001** (0.000)	0.000 (0.000)	0.000 (0.000)
Alcohol-related issues	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Maternity	0.005*** (0.001)	-0.001 (0.000)	-0.000 (0.000)
Emergency department	0.001 (0.002)	-0.000 (0.002)	-0.003 (0.002)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

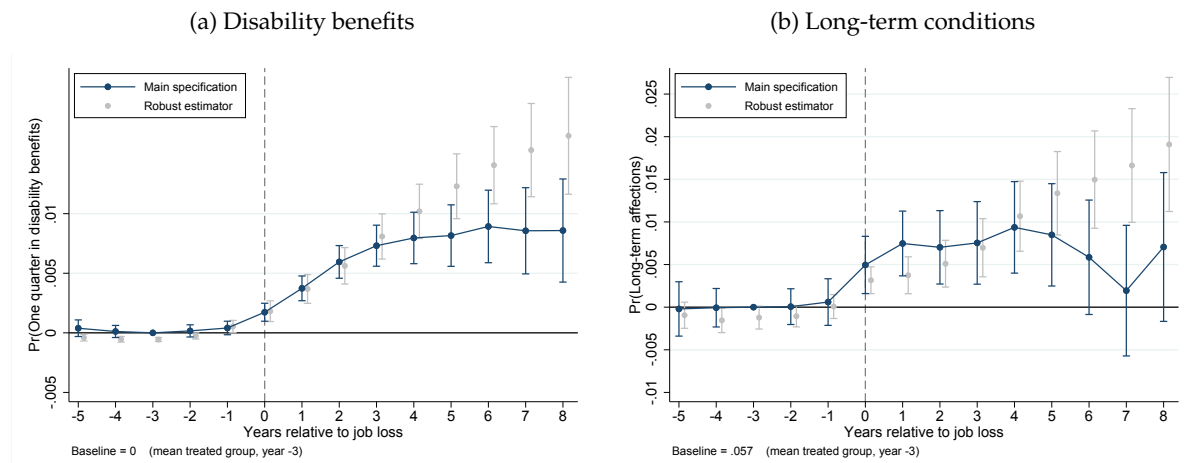
Figure 1.H.2: Heterogeneity (physical health)



Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 1.3). The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. The standard errors are clustered at the individual level.

1.H.2 Health Insurance

Figure 1.H.3: Effect of job loss on disability benefits and sick leaves



Notes: These figures plot the δ_k 's coefficients obtained from the staggered difference-in-difference regressions (equation 1.1) and the alternative estimator of Callaway and Sant'Anna [2021]. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -3) between the treated group and the control group. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Outcomes: (a) Probability of receiving disability benefits; (b) Probability of being recognised as suffering from a long-term condition. The standard errors are clustered at the individual level.

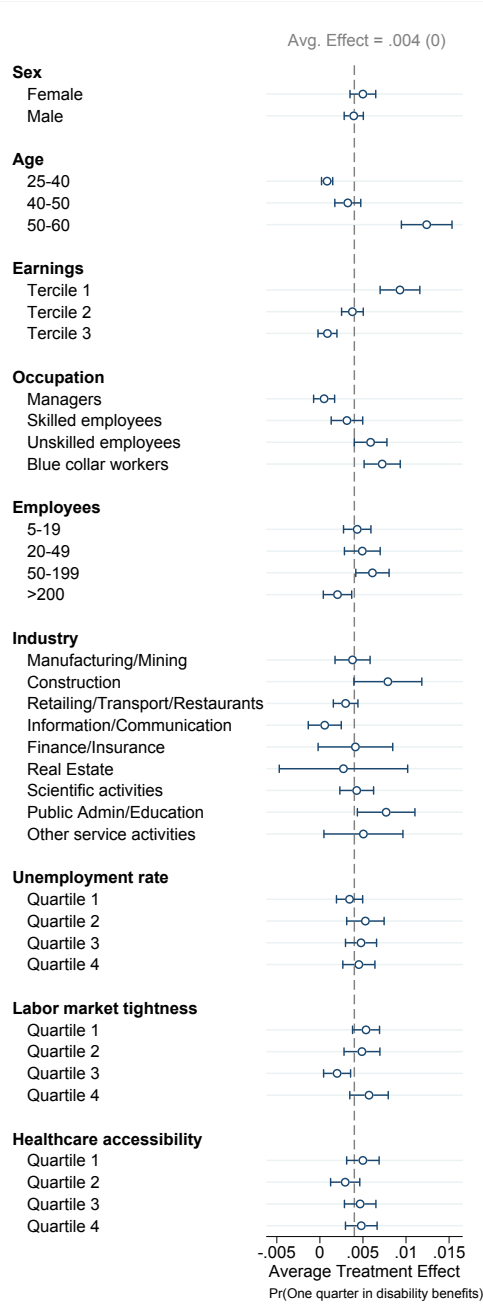
Table 1.H.3: Disaggregated outcomes for long-term conditions

	0-1 years	2-3 years	4+ years
Stroke	0.000606* (0.000277)	0.000641 (0.000402)	0.000513 (0.000511)
Spinal cord injuries	-0.0000177 (0.0000745)	0.0000110 (0.000129)	-0.000130 (0.000153)
Artheriopathies	0.000738* (0.000297)	0.000641 (0.000375)	0.000543 (0.000539)
Bilharziose	-0.00000177 (0.00000177)	-0.00000181 (0.00000182)	-0.0000340 (0.0000340)
Heart disease	0.0000251 (0.000360)	-0.000153 (0.000499)	0.000845 (0.000680)
Chronic liver disease	0.0000290 (0.000305)	0.000140 (0.000387)	0.000126 (0.000502)
Immune deficiency	0.000613* (0.000241)	0.000576* (0.000269)	0.000686* (0.000341)
Diabetes	0.000171 (0.000743)	0.000474 (0.000948)	0.00140 (0.00135)
Neurological disorders	0.000100 (0.000263)	0.000508 (0.000335)	0.0000500 (0.000467)
Haemoglobinopathies	-0.00000824 (0.0000139)	-0.00000848 (0.0000143)	-0.0000257 (0.0000464)
Haemophilia	-0.00000461 (0.000120)	0.0000206 (0.000155)	0.000318 (0.000219)
High blood pressure	0.000289 (0.000241)	0.000425 (0.000288)	0.000479 (0.000323)
Coronary artery disease	0.000266 (0.000480)	-0.000160 (0.000652)	-0.000259 (0.000919)
Respiratory insufficiency	0.000283 (0.000262)	0.000677 (0.000362)	0.000997* (0.000487)
Alzheimer disease	-0.000000370 (0.0000176)	-0.000000339 (0.0000853)	0.0000215 (0.000140)
Parkinson disease	0.000331* (0.000152)	0.000331 (0.000170)	0.000409 (0.000233)
Hereditary metabolic disease	-0.0000675 (0.000183)	0.0000241 (0.000220)	-0.000408 (0.000281)
Mucoviscidose	-4.32e-28 (5.52e-25)	-1.03e-27 (1.02e-24)	-8.70e-27 (1.44e-24)
Chronic kidney disease	-0.0000178 (0.000168)	0.000208 (0.000227)	0.000598 (0.000354)
Paraplegia	0.000121 (0.0000797)	0.000200* (0.000100)	0.0000814 (0.0000980)
Vasculities	0.000228 (0.000177)	0.000149 (0.000231)	-0.00000769 (0.000309)
Rheumatoid arthritis	-0.000124 (0.000323)	-0.00000286 (0.000404)	0.000514 (0.000515)
Psychiatric disorders	0.00227*** (0.000527)	0.00282*** (0.000734)	0.00259* (0.00101)
Crohn disease	-0.000127 (0.000296)	-0.000334 (0.000352)	-0.000187 (0.000390)
Multiple sclerosis	-0.000180 (0.000253)	-0.000326 (0.000297)	-0.000551 (0.000360)
Scoliosis	0.0000211 (0.0000381)	0.0000995 (0.0000636)	0.0000316 (0.000122)
Spondylitis	-0.000455 (0.000310)	-0.000268 (0.000410)	-0.0000850 (0.000482)
Following organ transplants	0.0000426 (0.0000421)	0.0000691 (0.0000534)	0.0001000 (0.0000675)
Tuberculosis	-0.000159 (0.000114)	-0.000159 (0.000108)	-0.0000311 (0.000157)
Malignant tumour	0.00229** (0.000806)	0.00210 (0.00110)	-0.0000262 (0.00153)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.H.4: Heterogeneity (health insurance)



Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 1.3). The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. The standard errors are clustered at the individual level.

1.I Details for the Channels section

This section provides more details on the individual labor market risk measured by the probability of finding a new job.

Methodology. First, I estimate re-employment probabilities using leave-one-out predictions. I train a logit model on the sample of treated individuals prior to displacement, based on demographic characteristics, employment history and health conditions (equation 1.I.3). Second I predict the employment probabilities for the treated individuals at the date of the event (equation 1.I.4). Then, I classify these individual labor market risks into terciles. Finally, I run three separate regressions from the main specification (equation 1.1) according to the tercile of the estimated individual risk.

$$P(\text{Employed}_{i,t+1} = 1 \mid t < E_i, D_i = 1) = F_\epsilon(\beta_0 + \beta_1 X_{1,it} + \beta_2 X_{2,it-1} + \beta_3 X_{3,it-2}) \quad \forall i \neq j \quad (1.I.3)$$

$$\hat{P}(\text{Employed}_{j,t+1} = 1 \mid t = E_j, D_j = 1) = F_\epsilon(\hat{\beta}_0 + \hat{\beta}_1 X_{1,jt} + \hat{\beta}_2 X_{2,jt-1} + \hat{\beta}_3 X_{3,jt-2}) \quad (1.I.4)$$

where $\text{Employed}_{i,t+1} = 1$ if individual i is employed in year $t + 1$;

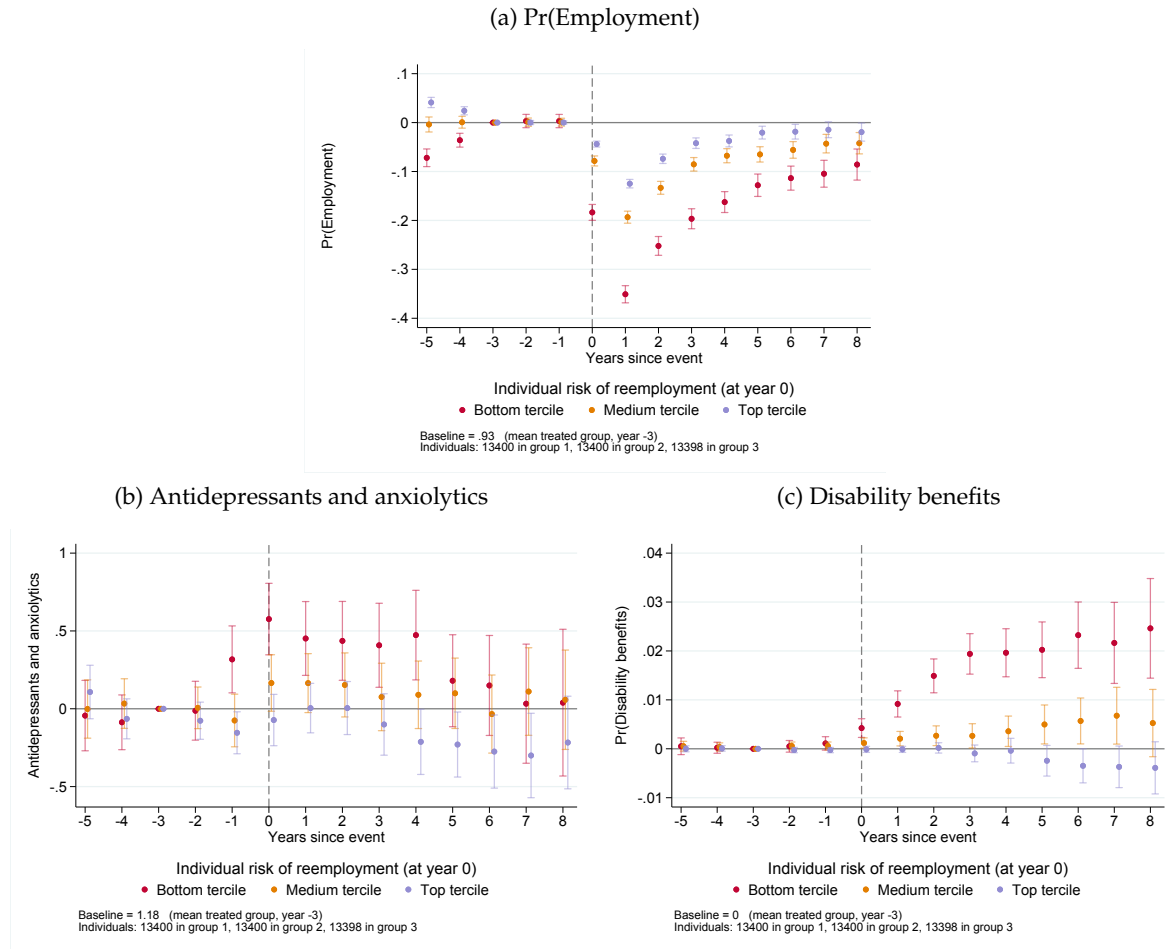
$X_{1,it}$ is a vector of characteristics of individual i at year t including sex, age, age squared, local unemployment rate, number of years in employment, department fixed effects, year fixed effects, last socio-professional category, last industry; $X_{2,it-1}$ is a vector of characteristics of individual i at year $t - 1$ including the employment status and ϵ follows a logit distribution; $X_{3,it-2}$ is a vector of characteristics of individual i at year $t - 2$ including the number of visits to GP, the number of hospital stay and the dummy for taking antidepressants and anxiolytics.

Validation. The estimated probabilities perform well in recovering the true employment status. Individuals with low estimated re-employment probabilities are less likely to be employed (Figure 1.I.1a).

Results. I then look at some health outcomes: antidepressants and anxiolytics and disability benefits. Workers with the lowest reemployment probabilities experience a higher increase in the use of mental health drugs and the effect is around 0 for those with the highest reemployment probability (however the effects are not significantly

different for each tercile) (Figure 1.I.1b). Workers who are less likely to be employed have a higher increase in the probability of receiving disability benefits (Figure 1.I.1c).

Figure 1.I.1: Predicted reemployment probabilities and health outcomes



Notes: These figures plot the δ_k 's coefficients obtained from the staggered difference-in-difference regressions (equation 1.1) on the three sub-populations defined by the predicted employment probabilities. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -3) between the treated group and the control group. The treated group is composed of all eligible individuals having experienced a plant closure. The control group is made of matched never-treated individuals. Year 0 denotes the year of the plant closure. Outcomes: (a) Probability to be in employment, (b) Number of boxes of antidepressants and anxiolytics, (c) Probability to receive disability benefits. The standard errors are clustered at the individual level.

Chapter 2

Automation at the Workplace: Implications for Workers' Health

This chapter is co-authored with Pauline Lesterquy.

Abstract

Technological progress has profoundly impacted employment, reshaping job content with varying effects across different skills and occupations. However, its overall impact on workers' well-being has received limited attention. This project addresses this gap by examining the effects of automation adoption at the firm level on workers' health, considering both physical and mental dimensions. By combining French individual healthcare records with firm-level data from customs and balance sheets in the manufacturing sector, we identify workers exposed to workplace automation. At the firm level, we find that productivity gains from automation outweigh displacement effects, as firms tend to increase hiring, particularly among blue-collar workers. At the worker level, automation has a limited impact on incumbent workers' earnings and employment status. However, it is associated with an increased use of antidepressants and anxiolytics, indicating a negative effect on mental health, while physical health outcomes remain unaffected on average. Additionally, older workers appear to be more vulnerable to the adverse effects of automation.

Keywords: Automation, Robots, Mental health, Physical health

JEL classification: I14, J24, 033

1 Introduction

Technological progress is constantly transforming the production process, redefining job tasks and reshaping labor market forces [Acemoglu and Restrepo, 2018; Autor, 2015; Brynjolfsson and McAfee, 2014; Frey and Osborne, 2017]. The most recent industrial revolutions, marked by the introduction of automation technologies and computers in the 1970s and followed by the advent of robots and digital technologies in the 2010s, have progressively transformed the nature of work. A significant portion of the workforce is exposed to these innovations: in OECD countries, up to 60% of jobs could potentially be automated in the future [World Bank, 2016].¹

The spread of automation technologies can significantly impact labor market outcomes by influencing labor productivity and shifting the demand for workers' skills. When these technologies act as substitutes for human skills, displacement risk increases, resulting in workers losing their jobs [Acemoglu and Autor, 2011; Acemoglu and Restrepo, 2018, 2020, 2022; Autor et al., 2003; Bessen et al., 2023; Goos, 2018; Graetz and Michaels, 2018]. Conversely, when automation complements human labor, it can lead to productivity gains [Acemoglu et al., 2020; Aghion et al., 2023, 2024; Bonfiglioli et al., 2024; Restrepo, 2023]. This process is a key factor behind job polarization, leading to changes in the workforce composition [Autor and Dorn, 2013; Braxton and Taska, 2023; Goldin and Katz, 1998; Goos and Manning, 2007; Lindner et al., 2022]. Additionally, it can result in the creation of new tasks where human labor holds a comparative advantage [Acemoglu and Restrepo, 2019].²

Automation may also have an impact on workers' well-being. The expected average effect on both physical and mental health remains ambiguous and deserves empirical investigation.³ First, workers may benefit from improved working conditions, fewer work-related diseases and lower workplace injury rates [Arntz et al., 2024; Gihleb et al., 2022; Gunadi and Ryu, 2021]. Alternatively, they might suffer from stress-related health issues or other negative effects from displacement risk, increased work intensity

¹A recent survey revealed that more than one-third of employees work at a pace monitored by computers in France [DARES, 2021].

²See Table 2.A.1 for a summary of the literature on technological change and labor market outcomes.

³See Figure 2.A.1 for a summary of the mechanisms.

or loss of meaning at work [Abeliansky et al., 2024; Bertoni et al., 2023; Blasco et al., 2022; Giuntella et al., 2023; Hernnäs, 2023; Lordan and Stringer, 2022]. These opposite mechanisms may hide substantial heterogeneity depending on worker characteristics [Arntz et al., 2024] and labor market institutions [Findeisen et al., 2024].⁴

In this project, we study how firm-level automation impacts workers' health outcomes by using administrative data from France between 2007 and 2019. We contribute in several ways to the understanding of the relationship between technological change and workers' health. First, we take advantage of a more direct measure of worker exposure than is typically used in the existing literature on automation and health, as we directly observe the automation investments made by the firms where the workers are employed.⁵ Second, we are able to compare the effects of different automation goods. Third, we investigate the channels of transmission of an automation shock to the different dimensions of workers' health, using very detailed administrative data on healthcare outcomes.⁶

Our empirical investigation relies on a new dataset that matches several administrative data sources, allowing us to thoroughly study the impact of technological adoption at the workplace on workers' health and economic outcomes in France between 2007 and 2019. In order to study workers' healthcare use and labor market trajectories, we rely on the representative panel of private sector employees sampled for the *Constances* epidemiological cohort. This allows us to build comprehensive worker-level indicators of physical health and mental health, based on administrative healthcare records (including visits to doctors, hospital stays, prescribed drugs and sick leaves). The richness of our health information allows us to break down and separately identify the effects on several dimensions, such as mental health and physical health conditions. It also contains panel information about individuals' careers (such as earnings, occupation, workplace identifier and social benefits). The panel dimension of our data proves useful both for identification and to pin down the dynamics of transmission

⁴See Table 2.A.2 for a summary of the literature on technological change and health outcomes.

⁵In the existing literature, many studies use a risk measure of exposure to automation, rather than a direct exposure, based on workers' characteristics [Blasco et al., 2022; Gihleb et al., 2022; Giuntella et al., 2023; Hernnäs, 2023] or at an aggregate sector or county level [Abeliansky et al., 2024; Gihleb et al., 2022; Gunadi and Ryu, 2021].

⁶Many existing studies rely on survey-based measures of health, with a limited number of outcomes and from a small number of waves.

through different mechanisms. Using the unique firm identifier of their main employers, we complement our dataset with firm-level administrative data. In order to detect firm-level automation, we leverage detailed product-level import flows from the Customs, balance sheet and income statement information, as well as firms' composition characteristics from matched employer-employee datasets. For now, we are focusing on automation in the manufacturing sector enabled by physical machines. Our first firm-level measure of automation is based on the balance-sheet net value of industrial equipment and machines [Aghion et al., 2024]. Our second measure is based on yearly firm-level imports of automation technologies and industrial robots. This is the most widely used measure in the recent literature on the economic effects of automation and robotization [Acemoglu and Restrepo, 2022; Bonfiglioli et al., 2024; Domini et al., 2022; Humlum, 2021]. One benefit is that it allows to draw upon an itemized list of machines and technologies using detailed product classification codes from the Combined Nomenclature. We focus on sizeable variations in our measures of firm-level automation investment, specifically those in the 90th percentile of the overall investment distribution, to characterize technological shifts initiated by firms.

First, we examine the effects that occur when a firm undergoes a technological shift. In a difference-in-differences setting, we compare firms that are heavily investing in automation to those that are making minimal automation investments.⁷ Our findings indicate that automating firms tend to increase their workforce, leading to compositional changes characterized by a higher proportion of blue-collar workers. This shift contributes to a reduction in the average hourly wage. Additionally, we observe an increase in labor productivity alongside a decline in the labor share. These results align with the positive perspective on automation, suggesting that in this context, the productivity benefits outweigh the displacement effects. Second, we analyze the impact of workplace automation on workers' labor and health outcomes. As working in an automating firm is not a random event, we combine a matching and difference-in-differences strategy. We implement a matching procedure to select a group of control workers, with a similar employment history and past health conditions, who are not

⁷Results are robust when using an alternative control of later treated firms.

experiencing an automation event the same year. We compare the evolution of health outcomes of exposed workers with that of a control group with a similar work history and healthcare use before workplace automation. Our findings show that incumbent workers, on average, do not experience significant changes in earnings. However, there is a clear age-related gradient: older workers face a higher risk of displacement, while younger workers see an increase in earnings. We also document a negative impact on mental health, with the use of antidepressants and anxiolytics rising by 27% on average. No significant changes are observed in physical health outcomes, except for a reduction in the use of anti-inflammatories among low-paid workers. Lastly, older workers and blue-collar workers significantly increase their take-up of disability benefits, and older workers spend an additional six days on sick leave.

In our next steps, we intend to implement complementary analyses and methodological improvements. First, we plan to instrument exogenous changes in the firm-level stock of automation capital to identify the causal effects of worker exposure to automation. Our two strategies would be i) to leverage a French reform that implemented a special amortization scheme for industrial robots and ii) to rely on a shift-share instrument based on productivity shocks of foreign suppliers of automation goods and their pre-existing relationships with French firms as in [Aghion et al. \[2024\]](#). Second, we will further investigate the mechanisms of transmission of automation to working conditions and workers' health by taking advantage of additional survey information available for a sizeable share of *Constances* workers, as well as complementary data from the *Survey on Working Conditions and Psycho-social Risks*. Lastly, we will include non-physical capital innovations, such as automation related to artificial intelligence, as they have very different consequences on workers' labor market outcomes [[Webb, 2020](#)]. To do so, we plan to use two firm surveys: the *Community Innovation Survey* and the *Survey on Information and Communication Technologies*.

The paper is organized as follows. Section 2 describes the data and samples. Section 3 provides some stylized facts on technological change and our definition of exposure. Section 4 describes the empirical strategy and presents the results for the firm-level analyses. Section 5 describes the empirical strategy and presents the results for the

worker-level analyses. Section 6 outlines the elements we would like to develop further in this project. Section 7 concludes.

2 Data

2.1 Worker-level data

We use the *Constances* data which is a French general-purpose epidemiological cohort. This cohort is designed to be representative of French private sector employees aged 18-69 years old at inclusion (see [Zins et al. \[2015\]](#) for details). A random sample of individuals was selected based on their age, gender and socio-economic status. Invitations were sent to the selected individuals and two populations were defined based on their participation: the active participants and the reference cohort. The *active participants* (around 170,000 individuals) voluntarily agreed to be included in this cohort: they complete self-administered questionnaires about their health, life events and occupation details (2012-2019; one questionnaire at inclusion and then annual follow-up questionnaires) and they are invited to attend one of the 22 Health Screening Centers (HSC) for a health examination.⁸ The *reference cohort* (around 300,000 individuals) is made up of individuals who were initially sampled but did not respond to the invitation. However, they are still passively tracked through their administrative records.⁹ For both the active participants and the reference cohort, healthcare consumption (2007-2019) and labor market history (1950-2019) are retrieved from administrative national databases.

Labor market history. Administrative data on the full labor market history of individuals are provided by the National Retirement Insurance Fund ("Caisse Nationale de l'Assurance Vieillesse", CNAV). For every year between 1950 and 2019, it provides occupation details from employers' reports (earnings, socio-professional category) and non-worked periods from social welfare organizations (unemployment, sick leave, maternity leave, disability benefits). The main variables for the analysis are defined as fol-

⁸Participants are invited to visit an HSC at inclusion and then every 5 years. Participants are randomly drawn from 19 targeted *départements* with a partner Health Screening Center (see map in Appendix 2.B.1). This is not a threat to representativeness because the structure of the population is essentially identical to that of France for the principal demographic, social, and occupational characteristics [[Zins et al., 2015](#)].

⁹This cohort was originally used to calculate adjustment weights for active participants (for representativeness purposes) and we received the CNIL authorization to include them in the study.

lows. We create a main annual employment status (employment, unemployment, sick leave and work accident, disability, maternity leave, retirement and inactivity) based on the number of quarters contributing to old-age pension (see Appendix 2.B.1.1 for details). We compute annual earnings, deflated to 2002 EUR, as the sum of labor earnings from the private sector (possibly from different employers).¹⁰ Earnings are thus equal to zero if an individual does not work in a given year. Finally, for each year, we can link the employee to his three main employers (workplace identifier SIRET), which is key for the identification strategy. The administrative labor market history is available both for the active participants and the reference cohort.¹¹

Healthcare consumption. Administrative data on the healthcare consumption of individuals are provided by the National Health Insurance Fund ("Système National des Données de Santé", SNDS). For every medical act between 2007 and 2019, it contains exhaustive individual data from different sources: reimbursement data (visits to doctors and other health professionals, prescribed drugs, medical devices), "long-term conditions" (serious chronic diseases exempt from co-payments), hospital discharge records (including diagnoses), medical and technical procedures. We consider three categories of healthcare consumption: visits to doctors, hospital stays and prescribed drugs. For each one, we can disentangle the medical care related to physical and mental health. For visits to doctors, we distinguish between general practitioners (GP hereafter), mental health specialists (psychiatrists and neuropsychiatrists) and other specialists. We identify hospital stays related to cancer, circulatory problems, alcohol-related issues and mental health based on the main diagnosis and we also include the use of the emergency department. We look at the consumption of specific drugs, namely antidepressants, anxiolytics and sleeping pills as well as painkillers, opioids and anti-inflammatories. The definition of healthcare outcomes is detailed in Table 2.B.2. The administrative healthcare consumption is available both for the active participants and the reference cohort.

¹⁰The sum also includes earnings of apprentices, domestic staff, au pairs or family workers, artists, authors and entertainment workers. There is a special case for people working in "Other Regimes" (e.g., civil servants, local authority and hospital staff): income is not provided but the individuals are in employment, so the value of earnings is left missing for this specific case. The old-age pension amount is also provided.

¹¹We plan to study labor market outcomes on a broader administrative dataset, see section 6.1.2 for discussion.

Sample. We define "eligible workers" as workers who have been employed at least once in an eligible firm (defined in section 2.2) between 2002 and 2019, and we only consider the years when they are within the working age range of 18 to 65 ($N = 49,023$).¹² Additionally, we follow workers even if they move across firms and industries during the period.

Matching with firm datasets. The *Constances* cohort provides information on individual healthcare consumption. For each worker and year, we observe the unique identifier of the establishment (and the firm) of the main employer. We match the *Constances* cohort with the firm datasets using this firm-level workplace identifier. This allows us to identify individuals who were exposed to the adoption of automation technologies in their firms.

2.2 Firm-level data

We also use firm-level data from 2002 to 2019 in France. We combine several datasets that cover the universe of French manufacturing firms.

Customs data. The datasets provided by the French customs provide exhaustive firm-level information on values and quantities of monthly imports by product and country of origin. We construct a firm-level proxy for automation, using the very detailed 8-digit product codes from the Combined Nomenclature (CN-8) to detect firm-level imports of "automation technology and robotics" products following [Acemoglu and Restrepo \[2022\]](#).

Balance sheet data. The *FICUS-FARE* dataset provides financial information from balance sheets and income statements of French firms. This dataset is nearly exhaustive since it is collected from mandatory tax reports. We use this source to extract the following firm-level variables: industry (defined as the main economic activity of the firm in the NAF rev. 2 classification), sales, value-added, and the stock of net tangible fixed assets. We also leverage detailed information from the raw tax files in the *Bénéfices Réels Normaux (BRN)* database to derive a second proxy of firm-level automation that is based on a direct measure of investment in "industrial equipment and machines".

¹²The age restriction is implemented to avoid mixing with the effects of retirement or old age on health.

Matched employer-employee data. The *Déclarations Annuelles des Données Sociales (DADS) - Postes* data provide comprehensive information at the job-spell level for each worker \times establishment \times year. This exhaustive dataset comes from employers' mandatory fiscal reports about their employees. For each year, we aggregate that information to create complementary firm-level variables (size, number of hours, occupational structure, wage bill, average hourly wage and share of female workers).¹³

Sample. We define the population of "eligible firms" as all importing firms in the manufacturing industry with at least 10 employees in each year from 2002 to 2019 (N = 21,167).^{14,15}

2.3 Proxies for automation in the firm

One of the main empirical challenges of this study is to characterize (variations in) exposure of workers to automation technologies in the firm. Ideally, we would like to observe the firm's production processes, measure their net stock and acquisitions of specific automation technologies and equipment in each period and observe which workers operate these machines. As these variables cannot be directly observed, we construct several alternative proxy measures for the purchases and stock of capital related to automation. Two proxies are commonly employed in the recent literature on the effects of firm-level automation, with a trade-off between more precise identification of automation products and complete coverage of firms' purchases.

2.3.1 Stock of industrial equipment and machines

Proxy 1 - Stock of industrial equipment and machines. Our first proxy of firm-level automation leverages the annual balance-sheet value and depreciation of "industrial equipment and machines" in the firm, following [Aghion et al. \[2024\]](#).¹⁶ This category of tangible fixed assets includes (i) industrial equipment, *i.e.*, "all equipment and machinery used for the extraction, processing, shaping and packaging of materials and

¹³All these variables are computed based on the workers employed in the firm as of December 31st.

¹⁴We exclude the "Installation and Repair of Machinery and Equipment" industry (5.8% of eligible firms) to avoid considering robot integrators or resellers, following [Bonfiglioli et al. \[2024\]](#).

¹⁵We define importing firms as the firms that have at least one import during the period 2002-2019, regardless of the nature of the imported products.

¹⁶It is reported in account number 215 in the tangible fixed assets, following French accounting standards.

supplies or for the provision of services" and (ii) industrial tools, *i.e.*, "instruments, tools and dies that are added to existing equipment to specialize it for a given task". It excludes tangible fixed assets related to land, buildings, transport, office and IT equipment and furniture.

Benefits and limitations. This method provides a broad-based and stable measure of the stock of industrial equipment over time. However, it is an aggregated proxy that lumps together several types of industrial fixed assets that may not all be related to automation.

2.3.2 Imports of automation technologies and industrial robots

Our second proxy takes advantage of the detailed customs data on the firm-level imports of automation products. The Customs collect information on all extra-EU imports and intra-EU purchases of goods, recorded at the firm-year-origin-product level using 8-digit product codes from the Combined Nomenclature. This allows us to identify firms that have imported automation goods in a given year by type of technology, as well as the value and origin of those imports.¹⁷

We will first examine the imports of automation goods based on a broad definition of "automation technologies" and then focus on "industrial robots" specifically.

Proxy 2a - Imports of automation technologies. Imports of automation goods are identified in the customs data based on the list provided by [Acemoglu and Restrepo \[2022\]](#).¹⁸ The list of automation products consists of the following categories in the HS6-2012 classification: industrial robots, dedicated machinery, numerically controlled machines, automatic machine tools, automatic welding machines, weaving and knitting machines, dedicated textile machines, automatic conveyors, and regulating and control instruments (\approx 500 different types of machines). Note that [Aghion et al. \[2024\]](#) extend that list to include more industrial machines in sectors other than the textile industry.¹⁹

¹⁷To study trends and variations in the net stock of imported automation and robotization goods (instead of flows), we compute two additional proxies based on an inventory approach using flows of past imports. The first proxy applies a 15% depreciation rate, while the second proxy uses linear depreciation over 8 years, following the standard French accounting practices for machinery. This allows for the computation of stock measures over the period from 2009 to 2019 (see Appendix 2.B.2.3 for details).

¹⁸The detailed list of codes in the HS6-2012 classification is provided in Table 2.B.3. More details are provided in Appendix 2.B.2.2 on the harmonization of the HS-2012 codes across the period.

¹⁹It should be noted that results are robust to using either of [Acemoglu and Restrepo \[2022\]](#)'s or [Aghion et al. \[2024\]](#)'s lists.

Proxy 2b - Imports of industrial robots. For the sake of robustness, we then consider a more restricted definition of automation goods, focusing on industrial robots.²⁰ These products are defined by the International Federation of Robots (IFR) as “automatically controlled, reprogrammable multipurpose manipulator, programmable in three or more axes, which can be either fixed in place or fixed to a mobile platform for use in automation applications in an industrial environment”. Given the very nature of the tasks that industrial robots are capable of performing, they are more likely to be pure substitutes for workers than other types of machines.

Benefits and limitations. This measure allows for a more precise identification of automation technologies, as it provides detailed information on the composition of the machines owned by the firms. In contrast with the balance-sheet measure that may be too broad, these import-based measures may suffer from opposite limitations. First, they are only available for the subset of importing firms. As such, we do not observe purchases from domestic suppliers, which means that we might be missing a portion of automation-related investment. However, this portion is likely to be quite small given that France does not hold a significant market share in the industrial robot industry.²¹ Second, firms are not required to report small intra-EU transactions if their annual total transactions are below a threshold.²² To circumvent this issue, we restrict the sample to large firms with more than 10 employees for which the threshold is less likely to be binding. Third, some firms importing automation products might be resellers. However, while purchases by robot integrators or resellers are possible, they are less likely in the manufacturing industry. As a precaution, we exclude the "Installation and Repair of Machinery and Equipment" industry, following [Bonfiglioli et al. \[2024\]](#).

²⁰ [Acemoglu et al. \[2020\]](#); [Bonfiglioli et al. \[2024\]](#); [Graetz and Michaels \[2018\]](#); [Humlum \[2021\]](#) also choose to focus on industrial robots, which are identified with HS-code 84795000 over the period.

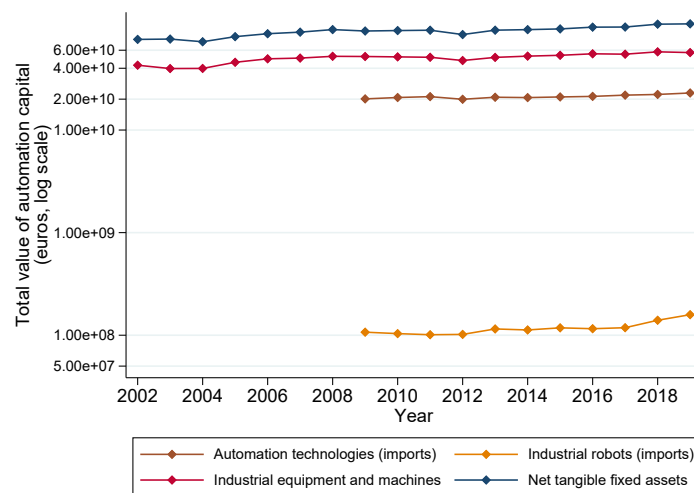
²¹ Using the share of global exports as a proxy for domestic production, [Bonfiglioli et al. \[2024\]](#) shows that France accounts for only 5% of the total volume of industrial robot exports, whereas Japan and Germany together account for about 50%.

²² The threshold above which firms have to report the CN8 code of traded products was 100,000 euros in 2002-2006, 150,000 euros in 2006-2010 and 460,000 euros in 2011-2019.

2.3.3 Aggregate trends and comparison of the proxies

The aggregate net balance-sheet value of tangible fixed assets increased before our period of interest between 2004 and 2008, decreased after the financial crises over 2008-2012, and increased again in the following years (Figure 2.1). The balance-sheet value of industrial equipment and machinery has evolved similarly and represents a sizable and stable portion of total net tangible fixed assets over the period 2009-2019. The measure of automation capital based on proxy 2a represents about half of the total value of industrial equipment, and it varies consistently with the industrial capital (proxy 1).²³ Finally, industrial robots appear to be a much less widespread type of automation goods, whose value is orders of magnitude smaller than total automation capital.

Figure 2.1: Evolution of the stock of tangible fixed assets and automation capital by proxy



Notes: This figure plots the total value of automation capital for our different proxies (in 2010 euros) owned by eligible firms from 2002 to 2019. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The stock of net tangible fixed assets and industrial equipment and machines are obtained from balance-sheet data. The proxy stocks of automation technologies (Acemoglu and Restrepo [2022]'s measure) and industrial robots are computed from the flows of imports using a linear depreciation over 8 years.

Overall, these aggregate measures do not indicate any trend that would point to automation-related structural change reflected in the total stock of capital. However, there can be significant variation in automation behavior at the firm level, which we will want to exploit to estimate our effects. We document some stylized facts with more detail in the next section.

²³However, we should keep in mind that this measure may capture other types of products due to harmonization so that we mostly focus on its variations for the purpose of analysis.

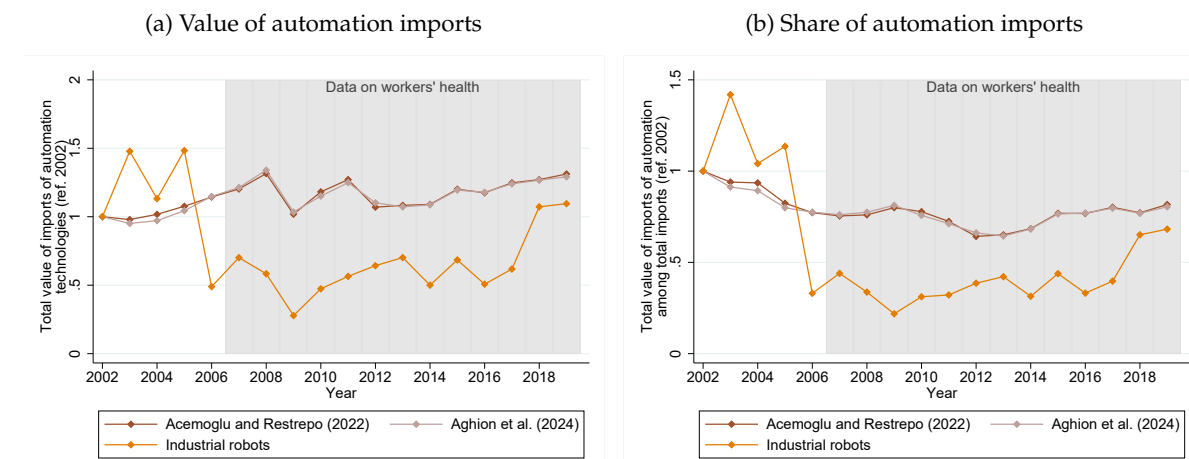
3 Stylized facts and treatment definition

3.1 Stylized facts

Fact 1. Aggregate imports of automation goods are quite stable over the period

Figure 2.1 displays macro trends in the imports of industrial robots and automation goods. The value of automation imports rose from 2002 to 2008, experienced a decline during the 2009 crisis, and then gradually returned to pre-crisis levels by 2019 (Figure 2.C.1a). The pattern is different for industrial robots, which saw a sharp decline in 2006, followed by a decrease in 2009 (Figure 2.1a). Imports and the stock of imported industrial robots also appear to have been increasing a bit more rapidly over the last few years 2018-2019; but these years are excluded from the reference period 2009-2016 used to define shocks. The proportion of automation imports within total imports has slightly decreased from an average of 4.3% in 2002 to 2.5% in 2012, before rising again thereafter. Overall, the share of automation goods and industrial robots in total imports remains quite stable over our analysis period 2007-2019 (Figures 2.1b and 2.C.1b).

Figure 2.1: Macro trends in automation imports



Notes: These figures plot the evolution of imports of automation goods between 2002 and 2019. Computations are made on the sample of eligible firms. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The light brown line corresponds to imports of products from the list in [Aghion et al. \[2024\]](#), the brown line to the list in [Acemoglu and Restrepo \[2022\]](#), and the orange line to industrial robots only. Outcomes: (a) Total value of imports of automation technologies, normalized to 2002 value; (b) Share of imports of automation technologies among total imports, normalized to 2002 value.

Fact 2. Concentration of automation imports over sectors remains relatively stable

The top five main suppliers of automation products to France represented around 60% of the total automation imports over the period 2002-2019. Germany and Italy have remained their largest suppliers over the period, with Germany holding a much larger share than other countries, albeit declining. China has been gaining market shares rapidly after its accession to the WTO and finally catching up with Italy around 2007. Imports of automation goods have also been concentrated in some industries. The three largest importing sectors, "computers and electronics", "motor vehicles" and "machinery and equipment", represented more than half of the total automation imports over the period (Figure 2.C.3).

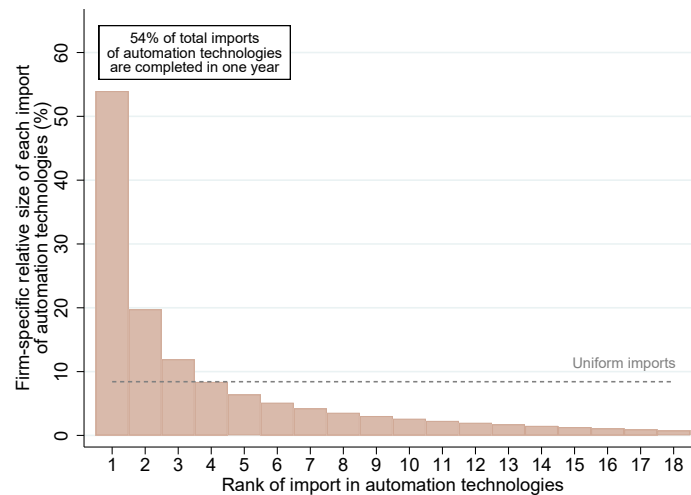
Although the stock of automation capital is also quite concentrated over a few manufacturing sectors, there are still around a dozen sectors holding a share of more than 2% (Figure 2.C.4). Structural change through sectoral reallocation has remained limited over the period 2009-2019. Indeed, the sectoral shares of industrial capital have remained quite stable. The top importer sectors remain globally the same for imported automation capital, but the shares are becoming a bit less concentrated in "computer and electronics" sector after 2014.

Fact 3. Firm-level adoption of automation technologies is lumpy

We explore the behavior of firms regarding investment in automation and robots over their lifecycle. Do firms smooth their investment in automation technologies, which would translate into imports of machines every year? Or do they follow infrequent but large movements, consistent with investment lumpiness as suggested by the recent literature in corporate finance? Figure 2.2 displays the share of annual automation imports (relative to total imports over the period) according to their rank in terms of value. Firm-level imports of automation goods in the firm are quite concentrated over time: on average, 54% of the firm's total automation imports are completed in a single year (Figure 2.2) and the number goes up to 85% when restricting the scope to only industrial robots (Figure 2.C.5). This investment behavior seems to be common across sectors, as is confirmed by the skewed distribution of firm-level imports by rank in all sectors (Figure 2.C.6). This supports the hypothesis that firms tend to initiate

technological changes all at once rather than gradually modifying their production structure. This also motivates our decision to use an event study approach, considering only one event per firm, which will be associated with its technological shift.

Figure 2.2: Lumpiness of imports of automation technologies



Notes: This graph illustrates the relative size of each import according to its rank. The y-axis displays the share of each import in the firm's total imports over the period; considering only imports of automation goods (Acemoglu and Restrepo [2022]'s measure). The x-axis represents the rank (according to its value in 2010 euros) of each import. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The dotted line represents the hypothetical scenario in which firms would import uniformly, *i.e.*, if imports were perfectly smoothed over time, throughout the entire period. Computations are based on a sample of eligible firms, from 2002 to 2019.

Fact 4. Workers' exposure to automation capital in the firm has remained stable

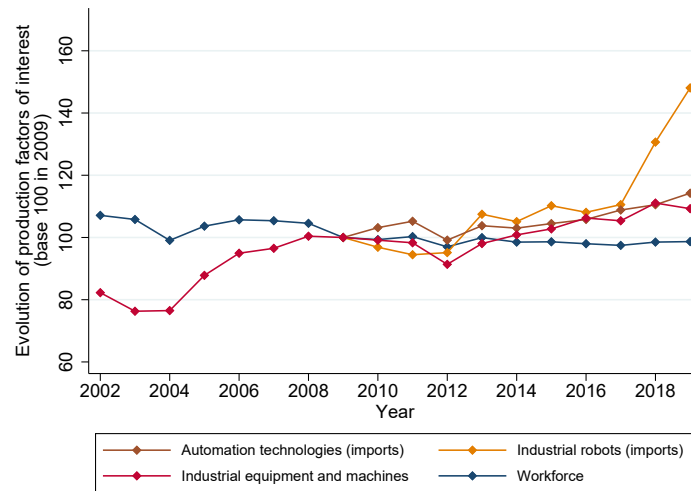
Over the period 2009-2019, the total number of employees, the stocks of industrial equipment and automation capital have remained relatively stable (Figure 2.3). These aggregate trends do not show an increase in the volume of machines per worker.

The share of workers who are exposed to a robotized (resp. automated) firm, defined as a firm that has a positive stock of robots (resp. automation technologies), is slightly increasing over the period (Figure 2.C.7a). More importantly, the use of automation technologies included in our broader measure is very widespread, considering that 98.6% of the firms in our sample own imported automation goods in 2018 (which weighs 99.7% of employment in the sample). Similarly, 98.7% of the firms own some industrial equipment (98.9% of employment). In contrast, only 5.8% of our firms own imported robots in 2018 (26.0% of employment). This suggests that, while we may identify new automation exposure during our period based on firm-level new adoption, most workers are already exposed to some extent to automation technologies and

industrial equipment at the beginning of the period, so we will use these measures to study variations in the "intensity" of exposure (intensive margin).

The aggregate value of automation stock per employed worker in our sample is stable over the period (around 30,000 euros₂₀₁₀, Figure 2.C.7b), similarly to the total value of automation capital and the total workforce employed in our sample of firms.

Figure 2.3: Evolution of the workforce and the stock of automation technologies



Notes: This graph plots the evolution of the workforce and the stock of automation capital by proxy, relative to their 2009 value. The net stock of automation capital is converted in constant 2010 euros using the deflator of gross fixed capital formation of machinery and equipment. The proxy stocks of automation technologies (Acemoglu and Restrepo [2022]'s measure) and industrial robots are computed from the flows of imports using a linear depreciation over 8 years. The total number of employees is computed from the matched employer-employee data. Computations are performed for the sample of eligible firms.

3.2 Treatment definition: characterizing a technological shift

We would like to identify a moment when firms undergo a technological shift. The underlying hypothesis is that this shift will go along with significant changes in production methods and, consequently, in the working conditions and tasks performed by workers. We attempt to detect these technological shifts using significant investments in automation-related capital (measured by the annual variation in the automation stock or intensity).²⁴

Our preferred strategy is to examine large investment events in the spirit of Aghion et al. [2024]. Previous literature has also used first investments or spikes in investment (see Appendix 2.D.1 for discussion). We select the largest firm-level investments based on their size relative to some pre-defined threshold (90th percentile in the distribution

²⁴However, it remains difficult to know if the purchases reflect the introduction of new technology in the firm or the replacement of existing equipment subject to capital depreciation.

of investments across all firms and years). Given the distribution of such investments in our sample, the 90th percentile seems more appropriate to define events in our case (Figure 2.D.1).²⁵

Event (1) - Proxy 1. An investment is computed as the log change in the net balance sheet value of industrial equipment and machines.²⁶ An automation event is defined as a firm making an investment in year t above the 90th percentile in the distribution of all possible positive changes.

$$\text{Event}_{jt} = \mathbb{1} \{ \ln(\text{Stock}_{jt}) - \ln(\text{Stock}_{jt-1}) \geq p90 \}$$

where Stock_{jt} is the net stock of industrial equipment and machines of firm j in year t and $p90$ is the 90th percentile in the distribution of all possible positive log changes across firms and years.

Events (2) and (3) - Proxies 2a and 2b. An investment is computed as the share of imports in automation technologies relative to total imports.^{27,28} An automation event is defined as a firm making an investment in year t above the 90th percentile in the distribution of all possible positive ratios across firms and years.

$$\text{Event}_{jt} = \mathbb{1} \left\{ \frac{AI_{jt}}{TI_{jt}} \geq p\tilde{90} \right\}$$

where AI_{jt} is the value of imports of automation technologies of firm j in year t , TI_{jt} is the total value of imports of firm j in year t and $p\tilde{90}$ is the 90th percentile in the distribution of all possible positive ratios across firms and years.

Alternative definition for industrial robots. The number of firms treated for industrial robot investment with this definition is very small because importing such products is rare (both at the aggregate and firm level). Thus, we define an additional treatment as the first import of industrial robots over the period, for robustness checks.

²⁵The 90th percentile is quite stable over time, our measure of treatment appears to capture automation events of comparable size and variation is not driven by a single year in the period (Figure 2.D.2).

²⁶To handle zeros, we use an inverse hyperbolic sine transformation. Formally, $\ln(\text{Stock}_{jt})$ in the above formula is in fact $\ln(\text{Stock}_{jt} + \sqrt{1 + \text{Stock}_{jt}^2})$. This approach allows for considering transitions from a zero stock to a positive stock (which might be over-represented) and not just transitions from one positive stock to another.

²⁷We do not use import flows directly because larger firms naturally have higher flows. Instead, we rescale them by total imports to obtain a relative measure.

²⁸We also used a similar measure to proxy 1 based on the proxy stocks computed from the flows of imports (following the procedure described in section 2.B.2.3) but the significant pre-trends do not allow us to conclude based on these graphs (see Figures 2.E.3a and 2.E.3b).

3.3 Samples of firms and workers

For the rest of the paper, we define the main automation event as a large investment in industrial equipment and machines (log change in the stock above the 90th percentage in the distribution, proxy 1).

3.3.1 Firms

For the firm-level analyses, we further narrow our sample to include only those firms that operated continuously from 2005 to 2019 and imported at least one automation good (as measured by [Acemoglu and Restrepo, 2022](#)) during the period 2002-2019 (N = 5,870).²⁹ We refer to this set of firms as the "narrowed sample". These restrictions help us select a more homogeneous group of comparable firms.³⁰

Treated group. A firm is treated at the first occurrence of an automation event. In the case of multiple automation events for a firm, we keep only the first occurrence, thus each firm has at most one automation event and is considered treated in all subsequent years. The treatment window is 2009-2016 so that we observe each firm for at least 2 years pre-treatment and 3 years post-treatment.³¹

Control group. Control firms are those that never experience an automation event. But they do automate on a smaller scale, having at least one instance of importing automation goods, as estimations are performed on the narrow sample of firms.

We identify 1,518 treated firms and 3,784 control firms (Table 2.1).³²

3.3.2 Workers

For the worker-level analyses, we focus on workers aged 20-60, working in eligible firms and with a strong labor market attachment: incumbent workers with a minimum of three years of tenure at an establishment at the time of the automation event and who

²⁹See section 2.C.5 for a comparison of the stylized facts between the eligible firms and the narrow sample.

³⁰Moreover, restricting the analysis to importing firms improves the credibility of imports being a good proxy for investment in automation capital, since firms already involved in trade may be more likely to source machines from abroad while we might think that other firms may source more machines from domestic firms or trade intermediaries (in which case, they do not provide a good control group when we use this proxy because of the unobserved domestic purchases of automation technologies). This is less of a concern when using balance-sheet measures of automation that are available for all firms.

³¹The firm is observed continuously from 2005 to 2019. We start considering automation events from 2009 and we drop any firm with an automation event after 2016.

³²See Table 2.D.1 for the sample size for the alternative treatment definitions.

were primarily working (referred to as "baseline restrictions").³³ We create a stacked dataset from different event cohorts c ($c \in [2009, 2016]$).

Treated group. A worker is treated during year c if he has a primary job in year $c - 1$ at a firm experiencing an automation event in year c (*i.e.*, between year $c - 1$ and year c).³⁴ Additionally, he has to meet the baseline restrictions in year $c - 1$. The criterion on tenure excludes new employees who might have been hired in anticipation of the upcoming technological shift to avoid selection bias. When workers are involved in multiple events, we only consider the first occurrence. The treatment window is 2009-2016 so each worker has at least 2 years of observation pre-treatment and 3 years post-treatment.³⁵

Control group. Working in an automating firm is not a random event. To address this concern, we select a group of control workers with similar employment and healthcare history in previous years. By performing a matching procedure, we ensure that the control group follows similar pre-treatment trends as the treated group. These trends serve as the counterfactual, representing what would have happened to the treated workers had they not undergone workplace automation.

For each cohort c , the pool of potential controls is made of all workers following the above baseline restrictions in year $c - 1$ and not experiencing an automation event in year c . Hence, control workers may either be treated later or never experience automation events.³⁶ A treated worker is matched in year c to an individual among the pool of potential controls with similar employment history and health conditions. We match on i) exact sex, 1-digit occupation and 2-digit industry and on ii) the propensity score (distance matching) where the predictors are age, earnings, tenure, firms' characteristics (size, growth, occupational structure, age), healthcare (visit to GP, hospitalization and prescribed drugs). All variables used for matching are measured in year $c - 1$,

³³The employment status restriction aims to exclude individuals who, while formally employed by the firm, spent the majority of the year in unemployment or on sick leave. Employment status is defined according to the annual classification detailed in the Appendix 2.B.1.1.

³⁴To increase the sample size, we also consider automating firms in the broader sample of eligible firms and not just those in the narrow sample.

³⁵The worker is observed continuously from 2007 to 2019. We start considering automation events starting in 2009 and we drop any workers experiencing an automation event after 2016.

³⁶Using only never-treated workers as the control group may be too restrictive because it involves conditioning on post-period treatment. We avoid that by only conditioning on pre-treatment variables and considering workers "not treated this year". For information, on average the matched control workers are made of 88% of never treated and 12% of later treated workers.

except the ones related to healthcare that are measured in year $c - 2$. We perform one-to-one nearest neighbor matching without replacement within each cohort c .

We match 1,465 treated workers to the same number of control workers in each stack (Table 2.1).³⁷

Table 2.1: Sample size - Firms & Workers

	Firms		Workers (<i>Constances</i>)	
	Treated	Control	Treated	Matched control
Investment: log change in the stock				
(1) Industrial equipment (proxy 1)	1,518	3,784	1,465	1,396
Investment: flows of imports				
(2) Automation tech. (proxy 2)	1,164	4,352	1,552	1,421
(3) Industrial robots (proxy 3)	16	5,747	.	.
Investment: first import				
(3) Industrial robots (proxy 3)	188	5,500	2,347	2,168

Notes: This table displays the size of the sample for each treatment definition. Columns 2 and 3 are the size of the treated and control groups respectively for firm-level analyses. Columns 4 and 5 are the size (*i.e.*, number of distinct individuals) of the treated and control groups respectively for worker-level analyses (after matching).

3.4 Descriptive statistics

Firms. Table 2.2 reports the mean and standard errors of the main outcomes variables for different subsets of firms for the first year of analysis (2007).

In the sample of all eligible firms, we document that adopting firms (*i.e.*, importing at least one automation good over the period) are larger and have higher sales on average than non-adopting firms (column 2). They also tend to have fewer blue-collar workers, higher labor productivity and hourly wages.

Turning to our narrow sample, the restrictions on the number of years of operation and importing decisions tend to select larger firms that are more capital-intensive with a lower labor share. They also tend to have higher hourly wages and labor productivity than the whole sample of eligible firms. In that narrow sample, treated firms (those in the top 10% of the distribution of log changes in either (1) the stock of industrial equipment, (2) the ratio of imports of automation goods, or (3) the ratio of industrial robots among total imports) are initially much smaller and have lower turnover than the non-treated firms. This suggests that our treatments are selecting smaller firms

³⁷We perform one-to-one matching without replacement for each cohort c , so a control worker may not be used multiple times within the same cohort but can be reused across different cohorts. They will appear with a unique individual-by-cohort identifier in each case. This explains why the total number of distinct control workers is slightly lower than the number of distinct treated individuals, although each cohort still has an equal number of treated and control workers. Moreover, a worker could be part of the control group for cohort c and part of the treated group for cohort c' with $c' > c$.

undergoing a large automation event akin to a technological shift in their production processes.

Comparing our treatment definitions, the treatment based on a broader industrial investment event (1) affects firms that are initially younger, smaller, have a higher share of female workers, a lower stock of industrial equipment, and lower imports of automation goods than firms subject to treatments (2) or (3). Additionally, we observe that the treatment (3) based on large investments in industrial robots seems to be undertaken by firms that have a very different profile than the average firm in our other two treatments. They tend to be larger, have a smaller share of blue-collar workers, higher turnover, a larger stock of industrial equipment, a lower labor share, and higher labor productivity and hourly wage compared to firms investing in the broader category of automation goods. This suggests possible heterogeneity in the behaviors and effects depending on the automation technologies.

These comparisons in average characteristics are quite similar at the time of event, when comparing firms undergoing an automation event to other firms in the same year. Compared to the narrow sample, firms undergoing a large automation event (1) or (2) tend to be smaller with lower labor productivity.

Table 2.2: Summary statistics (year 2007)

	Eligible firms		All	Narrowed sample		
	All	Never importing any autom. good		(1)	Treated firms (2)	(3)
Panel A: Employment						
Age of the firm	26.3 (18.8)	24.6 (17.1)	29.0 (19.6)	27.7 (18.6)	26.8 (16.9)	30.3 (16.0)
Size	134.1 (876.2)	35.3 (52.1)	187.3 (853.9)	63.6 (119.7)	87.5 (252.0)	141.4 (218.3)
Hourly wage	15.8 (4.1)	14.5 (3.7)	16.3 (4.0)	16.1 (4.2)	16.2 (3.8)	17.0 (2.8)
Share of blue collars	61.4 (20.4)	65.8 (19.4)	60.2 (20.0)	59.1 (22.1)	60.5 (21.8)	55.5 (20.9)
Share of females	29.3 (21.2)	29.5 (22.9)	28.7 (20.1)	28.6 (21.6)	24.3 (18.5)	25.3 (18.1)
Panel B: Financial variables						
Sales (k€)	43,191.0 (574,150.9)	6,362.9 (15,603.8)	68,229.3 (701,868.6)	16,052.2 (54,960.8)	17,135.6 (64,631.1)	27,078.8 (32,160.8)
Stock of industrial equipment (k€)	3,355.2 (49,159.9)	256.6 (930.0)	5,005.2 (53,218.0)	723.8 (4,479.8)	2,023.0 (13,095.1)	7,607.6 (21,579.2)
Imports in automation technologies (k€)	336.9 (4,278.9)	0.0 (0.0)	497.4 (5,545.1)	145.4 (1,032.9)	924.8 (1,1286.7)	356.7 (787.5)
Imports in industrial robots (k€)	2.1 (92.4)	0.0 (0.0)	4.5 (145.8)	1.9 (59.2)	5.9 (124.7)	138.8 (555.4)
Panel C: Factors						
Labor share	0.9 (7.4)	0.9 (1.7)	0.8 (1.4)	0.7 (0.4)	0.8 (1.1)	0.7 (0.2)
Labor productivity (k€\worker)	61.5 (67.9)	51.1 (36.0)	69.2 (90.5)	65.6 (97.0)	62.3 (27.0)	76.1 (36.2)
Observations	14,599	3,173	5,769	1,518	1,164	16

Notes: This table presents summary statistics of the main outcome variables for the year 2007, showing mean values with standard errors in parentheses. Columns 1 and 2 are based on the sample of eligible firms (importing manufacturing firms with at least ten employees). Column 1 provides statistics for all eligible firms, while Column 2 focuses on firms that did not import any automation goods in 2005-2019. Columns 4 through 7 are based on a narrowed sample of firms that operated continuously from 2005 to 2019 and imported at least one automation good. Column 4 includes all firms; Column 5 includes firms treated with event (1) (*i.e.*, firms in the top 10% of the distribution for the log change in the stock of industrial equipment); Column 6 includes firms treated with event (2) (*i.e.*, firms in the top 10% for the ratio of imports of automation goods among total imports); and Column 7 includes firms treated with event (3) (*i.e.*, firms in the top 10% for the ratio of imports of industrial robots among total imports).

Workers. Table 2.3 summarizes the average characteristics of workers in terms of demographics, employment history, and health outcomes, measured two years before the automation (or placebo) event. Prior to exposure to automation, treated workers are less likely to be female and tend to be slightly older (means statistically different, column 5). They also show better employment outcomes, including higher wages, longer job tenure, and employment in larger firms. This pattern suggests a selection process that favors more productive workers in firms adopting automation technologies, which tend to differ from non-adopting firms (as discussed above, Table 2.2). Treated workers are also more likely to be blue-collar employees and the proportion of white-

collar workers is similar across groups.³⁸ Despite these differences, health outcomes between treated and non-treated workers are relatively similar. Both groups have comparable mental health indicators (*e.g.*, psychiatrist visits, use of antidepressants and anxiolytics) and general health measures (*e.g.*, GP visits, hospitalization rates). However, treated workers are more likely to use painkillers than the random control group. The matching procedure successfully selected control individuals with similar labor market histories and healthcare usage prior to the automation event (few statistically significant differences, column 4).³⁹

The treated group consists of individuals who, on average, were 41 years old two years before the automation event, with 32% being female. These workers had been employed at their establishment for 10 years and earned approximately 29,000 euros annually. Of the treated workers, 54% are blue-collar and 20% are white-collar. They visited a GP an average of 3.6 times per year, 13.9% took antidepressants or anxiolytics, 61.7% used painkillers, and 11.2% had a hospital stay.

Workers may experience various transitions after the automation event, including staying at their current firm, moving to a different firm, or exiting the labor force. In the year following the automation event, 82% of treated workers remain employed at the same firm, 8% move to a different firm, 3% are unemployed, 3% are on sick leave, 3% have retired, and 0.5% receive disability benefits (Figure 2.D.3).

³⁸Other occupations, such as intermediate professions and clerks, are not displayed in the table.

³⁹One exception is that treated workers have slightly longer tenures and work in larger firms compared to matched control workers.

Table 2.3: Summary statistics (year $c - 2$)

	(1) Treated	(2) Matched controls	(3) Random controls	(4) p-value (1)-(2)	(5) p-value (1)-(3)
Panel A: Demographics					
Female (%)	32.08 (46.70)	32.08 (46.70)	33.42 (47.17)	1.000	0.281
Age	41.47 (9.72)	42.52 (9.76)	38.81 (12.90)	0.004	0.000
Panel B: Job characteristics					
Job tenure (years)	9.84 (7.11)	10.89 (7.81)	4.77 (6.01)	0.000	0.000
Annual earnings (2002-euros)	28,873 (17,640)	29,119 (17,694)	19,367 (21,639)	0.706	0.000
White collar (%)	20.14 (40.12)	19.25 (39.44)	19.50 (39.62)	0.546	0.551
Blue collar (%)	54.06 (49.85)	53.79 (49.87)	34.74 (47.62)	0.882	0.000
Workplace size	968.10 (2,508.69)	579.73 (1,918.19)	524.79 (1,520.79)	0.000	0.000
Panel C: Health outcomes					
Visits to GP	3.58 (3.41)	3.70 (3.77)	3.49 (3.72)	0.371	0.317
Visits to psychiatrist (%)	2.05 (14.17)	1.23 (11.02)	2.42 (15.37)	0.081	0.324
Antidepressants (%) and anxiolytics	13.86 (34.56)	14.40 (35.12)	14.18 (34.89)	0.671	0.724
Pain killers (%)	61.71 (48.63)	58.23 (49.34)	55.80 (49.66)	0.055	0.000
Hospital stays (%)	11.19 (31.54)	9.28 (29.03)	12.55 (33.13)	0.088	0.105
Observations	1,465	1,465	42,972	2,930	44,437

Notes: This table presents summary statistics two years before the (placebo) automation event for several sub-populations. The automation event is defined as a large investment in industrial equipment and machines (proxy 1). Column 1 is computed for treated workers; Column 2 for matched control workers and Column 3 for the pool of potential control workers (matched and unmatched control workers that satisfy the baseline restrictions). Note that the control workers in column 2 may appear multiple times if they were used as control workers for different stacks. Column 4 reports the p-value from the test for equality of means between column 1 and column 2. Column 5 reports the p-value from the test for equality of means between column 1 and column 3.

4 Firm-level results

We start by providing insights on what is happening at the firm level around large technological shifts (investment in industrial equipment and machines, proxy 1).

4.1 Specification

Difference-in-differences approach. We exploit variations in investments in automation-related capital across firms. Our strategy relies on comparing firms that invest heavily in automation goods to firms that automate only on a smaller scale, before and after

the automation event.⁴⁰ The baseline specification is a difference-in-differences event study design with a window of five leads and eight lags. The analysis covers the period from 2005 to 2019, with staggered events occurring between 2009 and 2016. We estimate the following equation on the sample of treated and never-treated firms (defined in section 3.3).⁴¹

$$\log(Y_{jt}) = \alpha_j + \lambda_{st} + \sum_{\substack{k=-5 \\ k \neq -1}}^8 \delta_k D_{jt}^k + \xi_{jt} \quad (2.1)$$

where Y_{jt} is the outcome of interest of firm j in year t . D_{jt}^k are distance-to-event dummies for firm j in year t (equal to 1 if year t is k years apart from the event year, and 0 otherwise), formally $D_{jt}^k = \mathbb{1}\{t = E_j + k\}$ with E_j the event year of firm j . α_j are firm fixed effects, controlling for firms' time-invariant characteristics. λ_{st} are 5-digit industry-by-year fixed effects, controlling for industry-specific time-varying shocks. ξ_{jt} is an error term. Standard errors are clustered at the firm level. The parameters of interest are the δ_k 's, which capture the association between investment in automation technologies and firm-level outcomes, *i.e.*, the evolution of the outcomes of treated firms k years after the event compared to the year before the event (δ_{-1} is normalized to zero), relative to the evolution of control firms.

Threats to identification. We consider these firm-level results as descriptive, given that the decision to adopt automation technologies might be endogenous.

A causal interpretation would require the assumption that treated and never-treated firms would have evolved similarly in the absence of treatment. There is concern that these two types of firms may not be comparable based on observed covariates, as never-treated firms are, on average, smaller, with a higher share of low-skilled workers and lower hourly wages (Table 2.2). Although we cannot formally test the parallel trends assumption, the absence of significant pre-trends is reassuring, as they do not show different growth in employment and sales before the investment event. In robustness analyses, we also test alternative specifications in which we use the timing of

⁴⁰Estimations are performed on the narrow sample of firms. Never-treated firms are those that never make significant investments in industrial equipment and machinery, though they do have at least one instance of importing automation goods.

⁴¹This specification is similar to [Acemoglu et al. \[2020\]](#); [Bonfiglioli et al. \[2024\]](#); [Domini et al. \[2022\]](#).

the event rather than the event itself, comparing firms that automate early to those that automate later. The results remain similar, both in sign and magnitude (see Appendix 2.E.3).

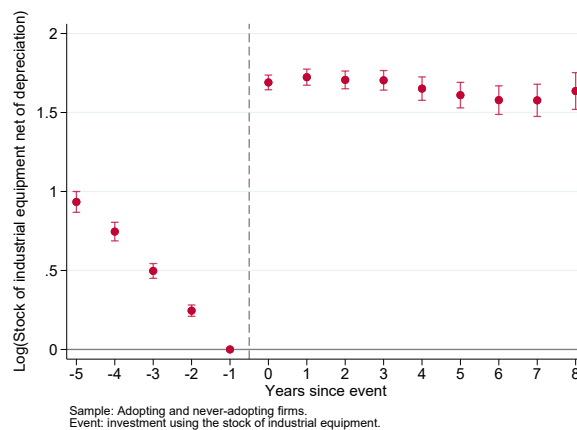
However, we cannot rule out the possibility of contemporaneous shocks. For instance, demand shocks may coincide in time with a firm's decision to invest in automation-related capital, potentially leading to both large automation shifts and more rapid firm expansion. This concern is common in the literature on technological change and some papers attempt to identify causal effects using shift-share IV designs [Aghion et al., 2024; Bonfiglioli et al., 2024; Koch et al., 2021]. Therefore, we will implement an IV identification strategy in further steps to estimate causal effects more robustly (see discussion on further steps in section 6.1).

4.2 Results

4.2.1 Validation exercise

First, we document what happens to the stock of industrial equipment and machines around the investment event. We observe that the stock starts decreasing before the technological shift and then quickly increases at a steady pace afterward (Figure 2.1).

Figure 2.1: Event studies at the firm level - results on automation



Notes: These figures plot the association of automation with different outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-differences regression of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment: Log change in the stock of industrial equipment and machines. Outcome: Log(stock of industrial equipment and machines). The standard errors are clustered at the firm level.

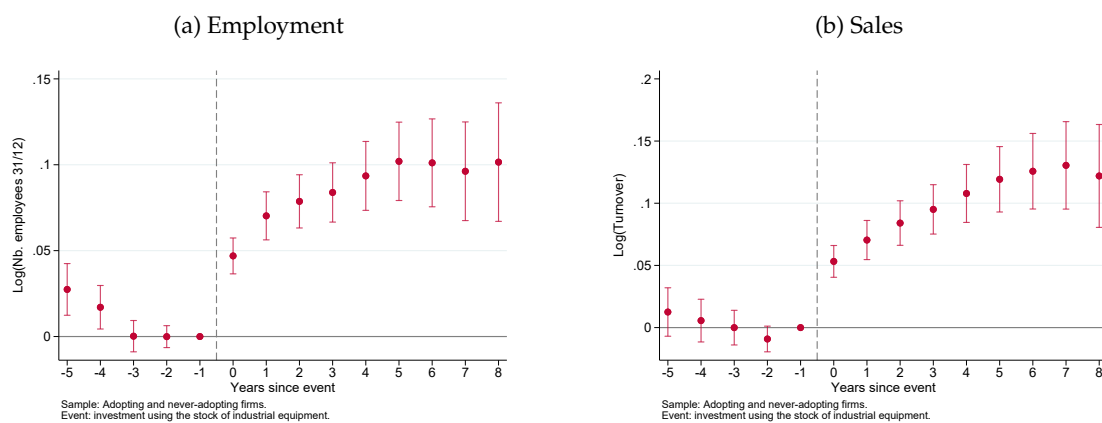
4.2.2 Main results

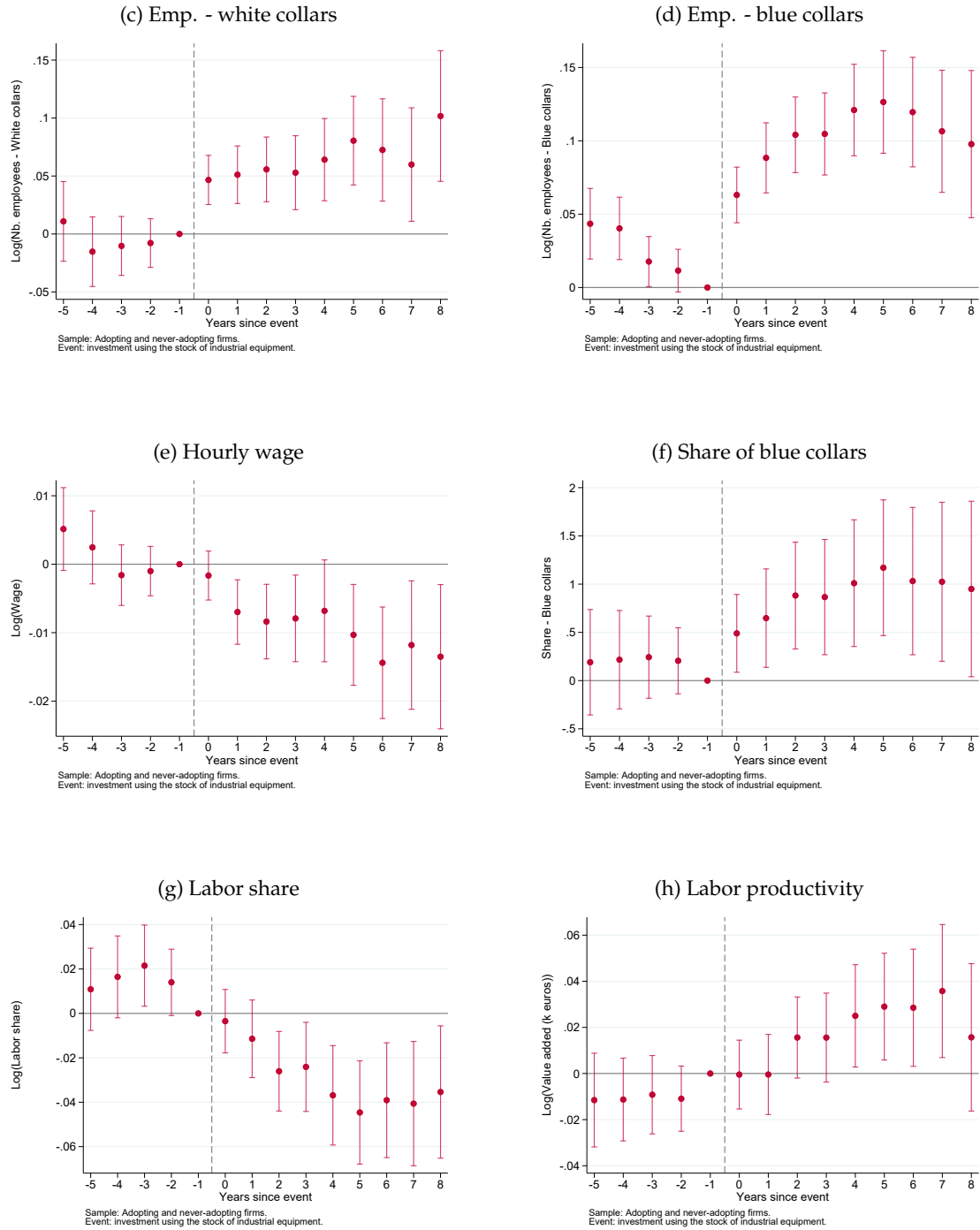
We consider the following outcomes: number of employees, sales, mean hourly wage overall and by occupation, share of blue-collar workers, labor share (as measured by the wage bill over the gross value added) and labor productivity (as measured by value added per worker).

We find that employment increases in firms that initiate a technological shift. The semi-elasticity of firm employment to the automation event is +0.08 on average over the 8 years following the automation event (Figure 2.2a). We also document a decrease in the labor share in value-added and an increase in labor productivity (Figures 2.2g and 2.2h). The sign of these effects is in line with other estimates from firm-level event studies in the literature. Employment is increasing for both white and blue-collar workers (Figures 2.2d and 2.2c) but we observe a significant increase in the share of blue-collar workers by almost 1 pp after the automation event (Figure 2.2f). The change in the workforce composition mostly explains the significant decrease in the mean hourly wage (Figure 2.2e), as we do not observe significant changes in hourly wages within occupations.

Overall, these findings suggest that the productivity channel outweighs the displacement effect at the firm level in our setting.

Figure 2.2: Event studies at the firm level - main results





Notes: These figures plot the association of automation with different outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-differences regression of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event (investment in industrial equipment and machines, proxy 1). The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Outcomes: (a) log(Number of employees); (b) log(Sales); (c) log(Number of white collar employees); (d) log(Number of blue-collar employees); (e) log(Mean hourly wage); (f) Share of blue-collar workers; (g) log(Labor share); (h) log(Labor productivity). The standard errors are clustered at the firm level.

4.3 Discussion

4.3.1 Robustness checks

Sample definition. Results are robust to a change in the definition of the sample. First, the results remain consistent when using an unbalanced sample of firms (without limiting to continuously operating firms), indicating that the findings are not driven by the entry or exit of firms. Second, the results are unchanged when we relax the assumption regarding the import of automation technologies, *i.e.*, by including firms that have never imported any automation goods in the control group.

Alternative definition of the event threshold. Results are robust to alternative thresholds for the event definition. For example, we define the treated group as firms in the 75th percentile of the distribution of log changes in the stock of industrial equipment and machines. This results in a less restrictive definition, thus including a larger number of firms in the treated group. We also find positive effects of automation on the number of employees within these firms. The effects are similar in magnitude but slightly higher, around 0.1 on average in the post-treatment period. Additionally, the growth rate is more consistently increasing under this specification, whereas it had become flatter in the main specification.

4.3.2 Changing automation event

We check whether these results are robust to alternative definitions of automation events, using imports of automation goods and industrial robots.

Validation exercise. Using the ratio of imports of automation goods (event (2)) or the ratio of imports of industrial robots (event (3)) to define the automation event, we do not observe any significant variation in the respective stocks before the event, but we document a significant increase in the stocks after the investment, which slightly decreases after five years (Figures 2.E.1a and 2.E.1b).⁴²

Results on employment. We then compare different automation events for the em-

⁴²The stock of automation goods follows a similar trend after the first import of industrial robots (Figure 2.E.6b). However, the trends when using the proxies for the stock of automation goods (Figure 2.E.1c) and industrial robots (Figure 2.E.1d) are more similar to the main event.

ployment outcome.⁴³ When considering alternative investment definitions (2) and (3) based on the ratio of imports of automation goods and industrial robots, we observe an increase in employment for the broad definition of automation goods (semi-elasticity is +0.04 on average across the eight years following the event, which is smaller than with the main event, Figure 2.E.3d) and no effect for industrial robots (Figure 2.E.3e). This is consistent with the definition of industrial robots, which perform tasks that are more substitutable than complementary to workers.^{44,45}

4.3.3 Changing specification

We then examine whether the results on employment are robust across different specifications. We use three alternative specifications: a staggered difference-in-differences design with only ever-treated workers (Figure 2.E.7b) and two stacked difference-in-differences designs with only ever-treated workers (Figure 2.E.7d includes distance-to-event coefficients for control workers while Figure 2.E.7c does not). Results are very consistent, both in sign and magnitude, across the main and the three alternative specifications.⁴⁶

5 Worker-level analysis

We now investigate the consequences of firms' adoption of automation technologies (investment in industrial equipment and machines, proxy 1) on workers' labor market and health outcomes (preliminary work).

⁴³The size of the treated and control groups for the alternative events are presented in Table 2.D.1. Figures 2.E.4 and 2.E.5 display the results for different outcomes: annual sales and the share of blue-collar workers respectively.

⁴⁴Figures 2.E.3b and 2.E.3c show the results for jumps in the stock of [Acemoglu and Restrepo \[2022\]](#)'s measure of automation goods and industrial robots, respectively. We only observe flows of imports so we compute proxies of the net stocks using equation 2.B.2. Results are similar whether using the formula with the proportional depreciation rate from equation 2.B.1 or the linear depreciation from equation 2.B.2 (Figure 2.E.2). Employment increases significantly after the jump in the stock of automation technologies, but significant pre-trends make the interpretation less convincing. Employment decreases significantly after the jump in the stock of industrial robots. We also find a decrease in employment after the first import of industrial robots (Figure 2.E.6b).

⁴⁵We also consider different measures of "intensity". When rescaling the change in the stock of industrial equipment and machines by the number of employees, to have a sense of how many workers are exposed to the machines, we also document that employment is significantly increasing (Figure 2.E.3f). Employment is significantly decreasing when considering a jump in the share of automation goods among physical capital or industrial equipment (Figures 2.E.3g and 2.E.3h). On the contrary, employment increases when there is a jump in the share of industrial equipment and machines among the physical capital (Figure 2.E.3i).

⁴⁶The alternative specifications are presented in Appendix 2.E.3. Results are also consistent using other outcomes, see Figure 2.E.8 for total turnover and Figure 2.E.9 for the share of blue-collar workers.

5.1 Specification

Difference-in-differences approach. We exploit variations in exposure to workplace automation across workers. Our strategy relies on comparing workers employed at firms undergoing an automation investment to workers at firms that do not experience such an event during the same period, before and after the automation event.⁴⁷ The baseline specification is a stacked difference-in-differences event study design with a window of five leads and six lags. The analysis covers the period from 2007 to 2019, with staggered events occurring between 2009 and 2016. We estimate the following equation on the sample of treated workers and matched control workers (defined in section 3.3).

$$Y_{it} = \alpha_i + \lambda_{st} + \sum_{\substack{k=\underline{T} \\ k \neq -1}}^{\bar{T}} \gamma_k D_{it}^k + \sum_{\substack{k=\underline{T} \\ k \neq -1}}^{\bar{T}} \delta_k D_{it}^k \cdot D_i + \epsilon_{it} \quad (2.2)$$

where Y_{it} are labor/health outcomes of individual-by-cohort i in year t . D_i is a dummy equal to 1 if the individual-by-cohort i is in the treated group. D_{it}^k are distance-to-event dummies (equal to 1 if year t is k years apart from the event year, and 0 otherwise), formally $D_{it}^k = \mathbb{1}\{t = E_i + k\} \forall t \in (\underline{T}, \bar{T})$, $D_{it}^{\bar{T}} = \mathbb{1}\{t \geq E_i + \bar{T}\}$ and $D_{it}^{\underline{T}} = \mathbb{1}\{t \leq E_i + \underline{T}\}$ where $\bar{T} = 7$, $\underline{T} = -6$ and E_i is the event year of individual-by-cohort i . α_i are individual-by-cohort fixed effects, controlling for individual-by-cohort time-invariant characteristics. λ_{st} are 2-digit industry-by-year fixed effects, controlling for industry-specific time-varying shocks. γ_k capture time-to-event effects for the control group (γ_{-1} is normalized to zero). This is necessary to control for the fact that both treated and control workers exhibit an upward work profile before the event due to the baseline restrictions. ϵ_{it} is an error term. Standard errors are clustered at the individual level. The parameters of interest are the δ_k 's, which capture the association between workplace investment in automation technologies and workers' outcomes, *i.e.*, the evolution of the outcomes of treated workers k years after the event compared to the year before the event (δ_{-1} is normalized to zero), relative to the evolution of control workers.

Heterogeneity analyses. Treatment effects may vary across different populations. To

⁴⁷Control workers are allowed to experience an automation event later.

explore this, we conduct heterogeneity analyses across several dimensions: demographics (such as gender and age) and job characteristics (including occupation, earnings, and establishment size). Take, for example, a heterogeneity dimension X with J groups (e.g., age). Each individual is assigned to a group j , based on their characteristics from the year prior to the automation event, which remains constant over time. To ensure clarity, we use a static two-way fixed effects model, where we pool together pre and post-event periods, and run separate regressions for each subpopulation $j \in 1, \dots, J$ based on the characteristic X .

$$Y_{it}^j = \alpha_i^j + \lambda_{st}^j + \gamma^j \mathbb{1}\{t \geq E_i\} + \delta^j \mathbb{1}\{t \geq E_i\} \cdot D_i + \epsilon_{it}^j \quad (2.3)$$

where Y_{it}^j is the health outcome at year t of individual i being in group j , $\mathbb{1}\{t \geq E_i\}$ is a dummy equal to 1 if period t is after the event. We plot the parameter of interest δ^j for each sub-population $j \in 1, \dots, J$.

Threats to identification. For reasons similar to those discussed in the firm section, we regard these results concerning workers as descriptive. To satisfy the parallel trend assumption, the labor market and health outcomes of treated workers should evolve similarly in the absence of the automation event. We do not find evidence contradicting this hypothesis in the pre-treatment period, largely due to the matching procedure which selects comparable workers. However, concerns about contemporaneous shocks remain, although they are less pronounced on the worker side. We are currently considering an instrumental variable (IV) identification strategy to provide causal estimates (see the discussion on further steps in section 6.1).

5.2 Results on Labor market outcomes

We begin by providing evidence on the long-term impact of automation on a range of economic outcomes (annual earnings and employment status) up to six years after the automation event.

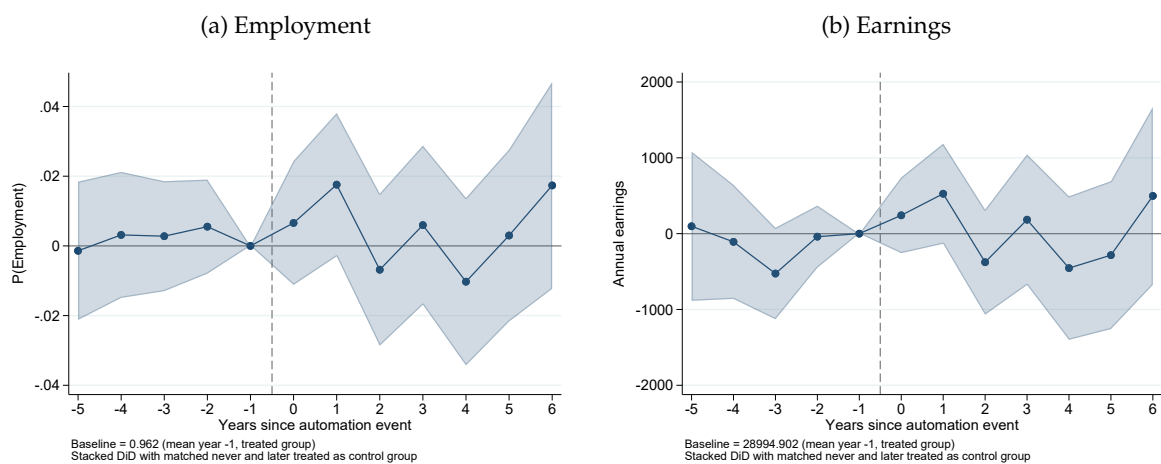
We compare the evolution of earnings for treated and (matched) control workers relative to year $c - 1$.⁴⁸ On average, we do not observe any significant changes in the

⁴⁸There are no restrictions on the post-treatment period for exposed workers. They can remain at the same firm, move to a

employment status of incumbent workers exposed to firm-level technological change (Figure 2.1a). Additionally, there are no observed changes at the intensive margin, as measured by total annual earnings (Figure 2.1b).

However, we observe heterogeneous patterns, especially based on age (Figure 2.F.1). Older workers (above 50) face a higher risk of displacement, with their probability of being employed decreasing by 6 percentage points and annual earnings falling by approximately 2,000 euros. In contrast, younger workers (aged 25–40) see a 3 percentage point increase in employment probability and an increase in annual earnings of about 1,800 euros. We do not find heterogeneous effects depending on occupation.

Figure 2.1: Event studies at the worker level - labor market outcomes



Notes: These figures plot the association of automation with workers' labor market outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) P(Employment); (b) Annual earnings. The standard errors are clustered at the worker level.

5.3 Results on Healthcare outcomes

We now investigate how workers' mental and physical health evolve after workplace automation. We consider different healthcare outcomes such as the use of prescribed drugs (antidepressants and anxiolytics, sleeping pills, pain killers and anti-inflammatories), visits to doctors (GP and psychiatrist), hospital stays (for any diagnosis, circulatory problems, cancer, mental health and alcohol-related issues), disability benefits and sick leaves.

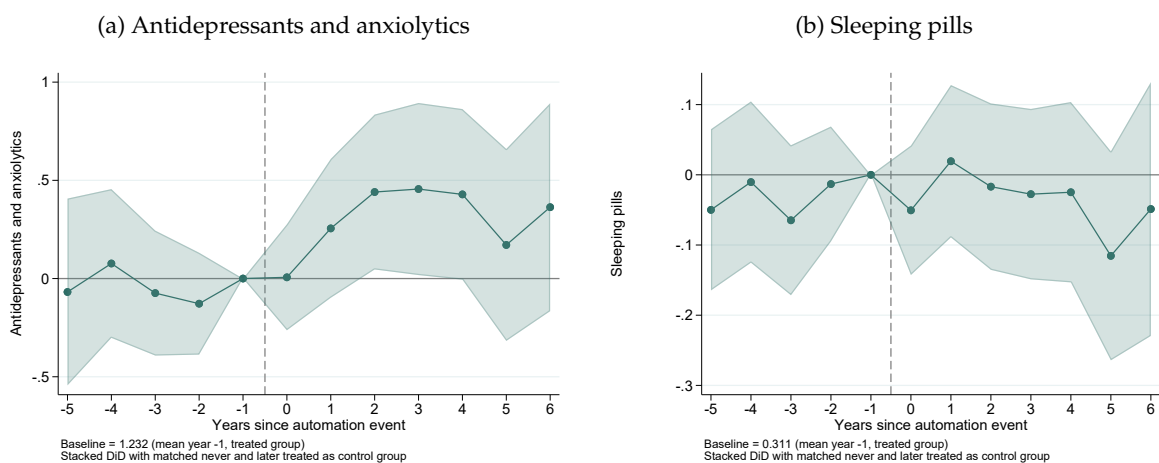
different firm, or leave employment (see Figure 2.D.3 for the employment status of the treated group following the automation event).

5.3.1 Mental health

We first look into mental health outcomes. The use of antidepressants and anxiolytics follows an increasing trend for exposed workers and is 0.33 boxes higher (27% increase relative to $c - 1$ mean) on average in the six years following the automation event (but only significant for the second and third years following the automation event, Figure 2.2a). We do not observe any significant difference on average in the use of sleeping pills (Figure 2.2b) nor in the number of visits to a psychiatrist or in hospital stays for mental health or alcohol-related issues (Figure 2.F.2).

The rise in the use of antidepressants and anxiolytics is primarily driven by women and older workers, both of whom increase their consumption by just under one box per year (Figure 2.4). Additionally, hospitalizations due to alcohol-related issues significantly increase among low-paid workers (Figure 2.F.4c). However, we do not observe any significant differences in other mental health outcomes across the other groups.

Figure 2.2: Event studies at the worker level - mental health outcomes



Notes: These figures plot the association of automation with workers' mental health outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) Number of boxes of antidepressants and anxiolytics; (b) Number of boxes of sleeping pills. The standard errors are clustered at the worker level.

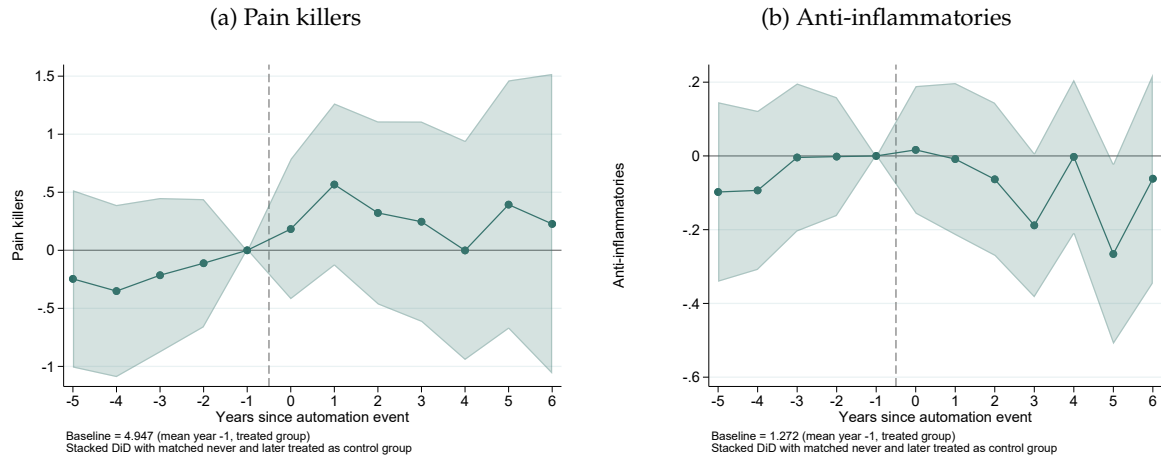
5.3.2 Physical health

We now turn our attention to general or physical health outcomes. On average, we do not observe any significant changes in the use of painkillers (Figure 2.3a), nor in the number of visits to general practitioners (GPs) or hospital stays for various diagnoses

(Figure 2.F.3). However, there appears to be a decreasing trend in the use of anti-inflammatories (Figure 2.3b).

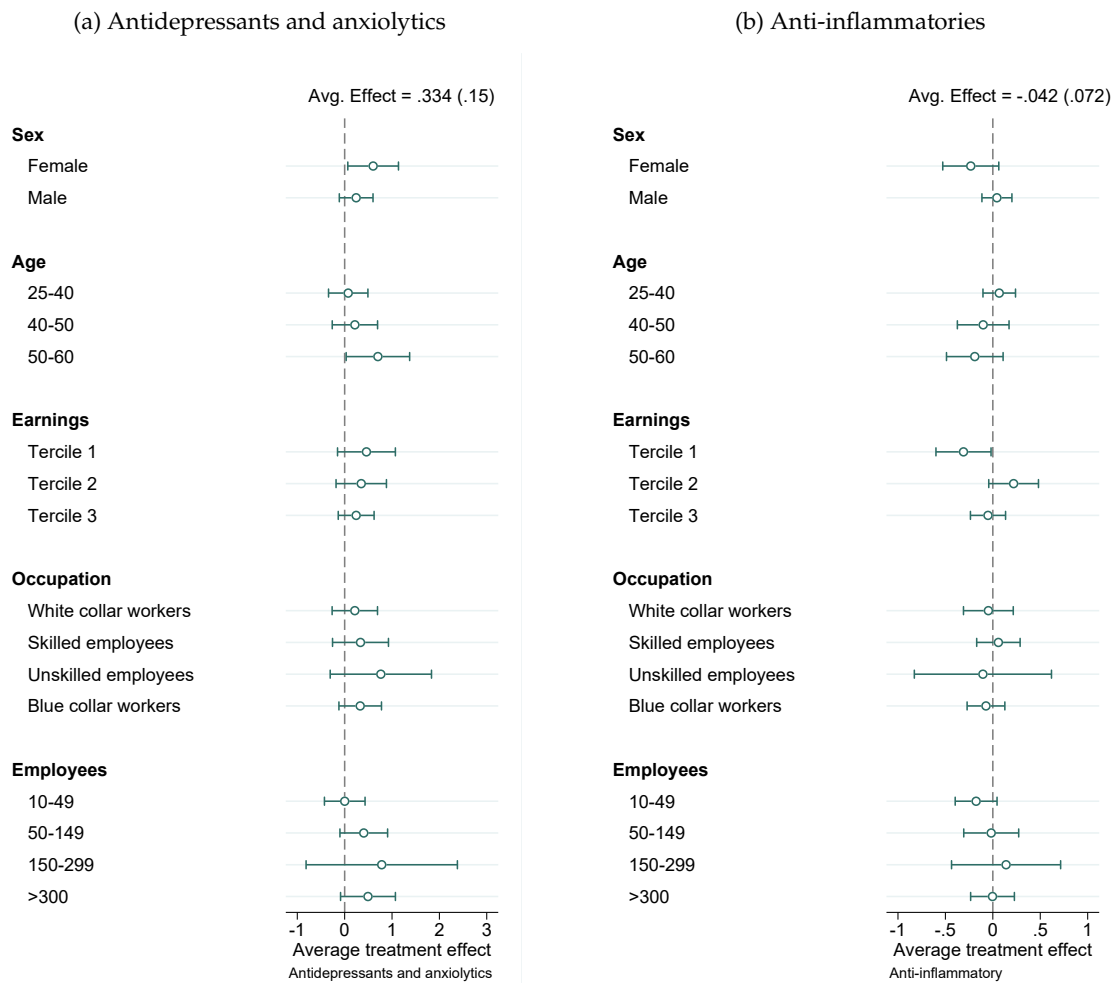
The number of visits to GPs increases among older workers (Figure 2.F.4c). The decline in the use of anti-inflammatories is primarily driven by low-paid workers, who show a decrease of half a box per year (Figure 2.4).

Figure 2.3: Event studies at the worker level - physical health outcomes



Notes: These figures plot the association of automation with workers' mental health outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) Number of boxes of pain killers; (b) Number of boxes of anti-inflammatories. The standard errors are clustered at the worker level.

Figure 2.4: Heterogeneity



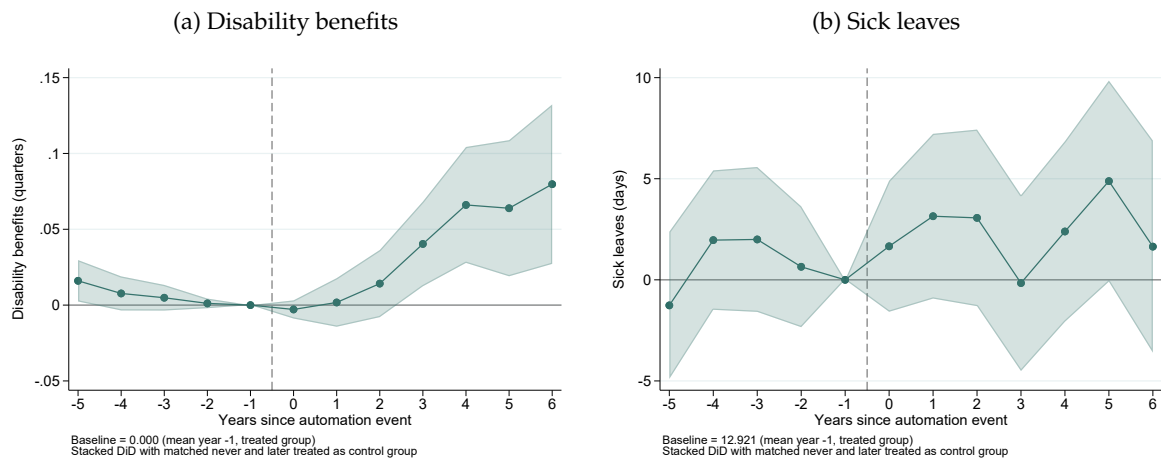
Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 2.3). δ^j measures the difference in the evolution of the outcome (between the pre and post-period) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Outcomes: (a) Number of boxes of antidepressants and anxiolytics; (b) Number of boxes of anti-inflammatories. The standard errors are clustered at the worker level.

5.3.3 Health insurance

Finally, we examine two health insurance outcomes: disability benefits and sick leave. Exposed workers show a significant and persistent increase in their uptake of disability benefits. We do not observe any significant differences in the number of days in sick leave on average (Figure 2.5b).

The uptake of disability benefits rises significantly among older workers, males, low-paid workers, and blue-collar employees (Figure 2.F.6). Notably, older workers experience an increase of 6 days per year in sick leave, despite being less likely to be employed.

Figure 2.5: Event studies at the worker level - health insurance



Notes: These figures plot the association of automation with workers' mental health outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) Number of quarters in disability benefits; (b) Number of days in sick leave. The standard errors are clustered at the worker level.

5.4 Discussion

The above sections present a first set of preliminary results on the association between workplace automation and workers' labor market and health outcomes. We aim to complement these analyses with robustness tests and by examining other measures of exposure.

Sample size. The current size of the treated group is quite small, and we will attempt to increase it by using less discriminatory thresholds (such as the 50th or 75th percentiles in the distribution of investments) and by imposing fewer restrictions on the selection of eligible firms.

Treatment definition. We also plan to explore an alternative treatment measure to capture more precise worker-level exposure. For example, we consider the variation in the stock of automation goods, adjusted for the number of workers or blue-collar workers, to assess the intensity of automation and the extent to which workers interact with machines. The effects observed at the firm level remain consistent with this approach (Figure 2.E.3f).

Specification. We also plan to perform robustness checks by comparing other specifications such as changing the definition of the control group to those employed at

later treated firms (following Bessen et al. [2023]). We will also implement recent difference-in-differences estimators [Borusyak et al., 2022; Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfoeuille, 2020b; Sun and Abraham, 2021]. Recent literature on difference-in-differences with staggered treatment has shown that this standard two-way fixed effects approach may be biased in the presence of heterogeneous treatment effects across time and between groups [de Chaisemartin and D'Haultfoeuille, 2020a; Goodman-Bacon, 2021]. Automation-related treatment effects are likely to be heterogeneous over time (as types of machines evolve) and between individuals (due to different exposures based on occupation). However, these complications might be mitigated in our current stacked design, which uses clean controls (matched treatment-control groups for each cohort, where controls share similar work histories but are not treated in the same year), thus limiting the risk of forbidden comparisons.

6 Further steps

We plan to further develop this paper in two ways: first, by improving the identification strategy to obtain causal estimates, and second, by incorporating additional datasets to capture measures of automation related to non-physical assets.

6.1 Improving the identification strategy to estimate causal effects

First, we plan to instrument exogenous changes in workers' exposure to automation in their workplace to estimate the causal effects on worker-level health outcomes.

6.1.1 Find an instrumental approach

- **The Action Plan for Firms' Growth and Transformation (PACTE) reform in 2013.**

Our preferred identification strategy to estimate causal effects would be to leverage a French reform that implemented an accelerated amortization scheme for SMEs' investments in industrial robots from October 2013 to 31st December 2016, introducing an incentive to invest through an exogenous decrease in the cost of robots for these firms. Moreover, the firms had to report the amortization in their detailed tax returns, so that we are able to measure these expenses using the most

detailed *Bénéfices industriels et commerciaux (BIC)* dataset on yearly firms' tax returns. An advantage of this approach is that we directly observe an exogenous increase in the automation capital of the firm. However, the population of beneficiary firms is limited, so we need to check whether we have a sufficient sample size of exposed workers. This strategy would provide clean estimates of causal effects on workers' health.⁴⁹

- **Shift-share instrument:** An alternative way to recover causal estimates would be to construct Bartik-style instruments in a shift-share IV design akin to [Aghion et al. \[2024\]](#). The *shocks* can be constructed from changes in market shares of suppliers (by origin and HS6 product) in countries similar to France (such as other EU countries and Switzerland). The shocks reflect variation in the costs of imports of automation goods that is mostly exogenous to the decisions of importing French firms (*i.e.*, assuming that the decisions of French importers do not drive the changes in unobserved quality-adjusted prices of machines). The *shares* should reflect the pre-determined exposure of firms to each supplier country by HS6 product. [Aghion et al. \[2024\]](#) construct them using the shares of each supplier in each HS6 product, based on the firm's imports in the reference period 0 before the shocks. The strength of these instruments requires that the shares reflect the actual exposure of firms to productivity shocks at the HS6-supplier level, which would be justified for example by the existence of switching costs making the buyer-supplier relationship sticky. If we pick this approach, we will attempt to refine the instruments to take into account our empirical finding that automation investments are often lumpy at the firm level.⁵⁰ Considering that we are specifically interested in estimating worker-level effects of automation, we will need to construct worker-level exposure *shares*. The first solution is to use the firm-level

⁴⁹Estimates will be interpreted with caution, as there may be concerns about external validity given that this estimation strategy relies solely on variation in small firms. These concerns may be alleviated by event studies and further robustness checks.

⁵⁰[Bonfiglioli et al. \[2024\]](#) propose an alternative strategy to filter out demand shocks and identify exogenous changes in robot adoption happening for technological reasons. Since most of the variation is across firms, they estimate a long-difference specification of their model and construct instruments that reflect differences in firms' pre-determined propensity to automate within an industry. This firm-level instrument, quite comparable to *exposure shares* from the shift-share strategy, is constructed as the interaction of (i) an industry-level measure of the initial average robot intensity of capital of other firms in the industry and (ii) a firm-level measure of replaceability based on the occupational composition of the firm's workforce and occupation-based replaceability scores derived from [Graetz and Michaels \[2018\]](#). Therefore, the construction of the shares relies more on the assumption that level differences in initial industry-level capital and firm-level workforce compositions predict differences in sunk costs and frictions in automating production processes, rather than assumptions on trade relationship stickiness.

shares as a measure of exposure for the workers employed in that firm. However, this might make the instrument weaker. Alternative instruments could be constructed with direct worker-level measures of exposure, by exploiting other sources of data (see section 6.1.2).

6.1.2 Refine the worker-level identification and analysis

Ideally, we would like to observe exposure to automation at the worker level. As this is not observed, we currently consider all workers in a treated firm as treated to estimate our effects. To have more detailed worker-level information on occupation and labor market outcomes than in the *Constances* panel, we may use another complementary data source to construct the worker-level measures of exposure:

- **Occupation-based approach.** The DADS panel is an administrative database of matched employer-employee information collected by the French statistical institute INSEE. It covers 1/12th of the workforce, including private sector employees, hospital public service workers, local government employees and state agents. This dataset contains very detailed data on working hours, wages and occupations. This would allow us to improve our identification strategy, using occupation-level scores of *automatability* from the literature to discriminate between occupations that are more at risk of automation or more complementary to these technologies. We will assess whether the new findings on labor market outcomes are consistent with those from the epidemiological cohort and complement our understanding of mechanisms. As a complementary source, it will also allow for a more precise analysis of employment outcomes and transitions on a larger sample size.
- **Task-based approach.** Worker-level exposure to *automation risk* may be better measured at the task level rather than at the occupation level. Depending on the size of the sample, we might be able to construct such worker-level measures of exposure using INSEE's *Enquête Conditions de Travail* survey waves of 2013 and 2016.

6.2 Decomposing between different types of automation technologies

So far, this project has primarily focused on physical capital, such as industrial machines, automation technologies, and robots, that we identify through customs and balance sheet data. We intend to broaden our scope to encompass artificial intelligence and software, as these may have significant and distinct effects on workers' health. To do so, we plan to use surveys on information and communication technologies (ICT) and data from the Community Innovation Survey (CIS). This approach will enable us to differentiate between various types of automation products such as machines, computerization, and digitalization technologies.

7 Conclusion

This paper investigates the impact of workplace automation on workers' health in the manufacturing sector in France from 2007 to 2019. While previous studies have documented changes in tasks driven by skill-biased technological change, our goal is to understand the consequences for the physical and mental health of exposed workers. We detect the penetration of automation at the firm level by leveraging rich administrative data from customs and balance sheets. We define two proxies for automation investment based on the stock of industrial equipment and machines, and the imports of automation goods and industrial robots. We begin by describing the impact of automation events at the firm level and show that productivity gains from automation outweigh displacement effects. In a difference-in-differences setting, we compare firms making significant investments in automation with those that implement only minimal automation. Automating firms tend to increase their number of employees, especially blue-collar workers. We then rely on detailed individual healthcare data matched with firm data to identify workers exposed to automation investments. In a difference-in-differences setting, we compare workers employed at firms undergoing an automation investment to those at firms that did not experience such an event during the same period. We provide preliminary insights into the impact on workers' labor market and health outcomes. Younger workers tend to benefit from higher wages,

while older workers face increased displacement risks and are more reliant on disability benefits and sick leave. Mental health outcomes worsen on average, as reflected in the increased use of antidepressants and anxiolytics. However, physical health impacts appear limited, with only a reduction in the use of anti-inflammatory medication among low-paid workers. In the next steps, we plan to explore the mechanisms at play and improve the identification strategy to retrieve causal effects.

Appendices to Chapter 2

2.A Motivation

2.A.1 Mechanisms

Figure 2.A.1: Mechanisms

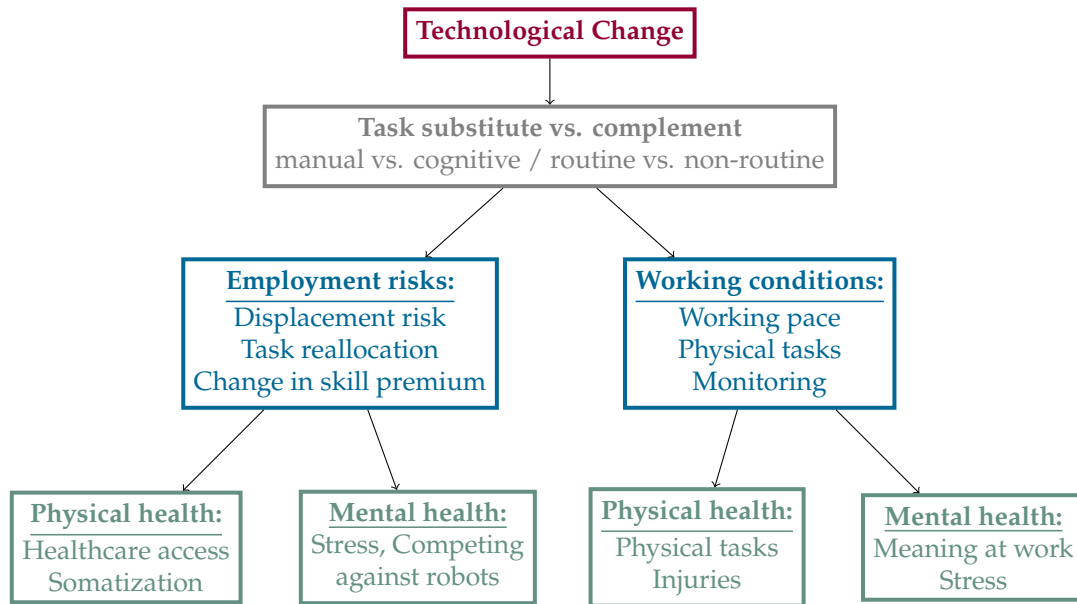


Table 2.A.1: Literature review on technological change & labor market outcomes

Paper	Context	Level	Technology	Shock	Causality	Results (employment)
Acemoglu et al. [2020]	France 2010-2015	Local labor markets	Industrial robots (imports + other sources)	Any adoption		Positive
Acemoglu and Restrepo [2020]	USA 1990-2007	Local labor markets	Industrial robots (IFR)	Industry-level exposure		Negative
Aghion et al. [2024]	France 1995-2017	Firms & Industries	Industrial equipment & Automation technologies (imports, Acemoglu and Restrepo [2022])	Investment event Treated firms = top 10%	Shift-share IV Productivity shocks in suppliers Pre-existing trade shares	Positive
Bessen et al. [2023]	Netherlands 2000-2016	Firms & Workers	Automation costs	Spike	Workers only	Negative
Bonfiglioli et al. [2024]	France 1994-2013	Firms	Industrial robots (imports)	First adoption	Robot exposure instrument (tech. characteristics)	Positive correlation but negative (causal)
Domini et al. [2022]	France 2002-2017	Firms	Automation [Acemoglu and Restrepo, 2022] & AI-related goods	Largest import		Positive. No effect on within-firm wage inequality.
Graetz and Michaels [2018]	17 countries 1993-2007	Industry × Country	Industrial robots (IFR)	Industry-level exposure		Small. Negative for low-skilled workers.
Humlum [2021]	Denmark 1995-2015	Firms	Industrial robots (imports) + Survey	Any adoption		Positive (tech workers), negative (prod. workers)
Koch et al. [2021]	Spain 1990-2016	Firms	Robots	Any adoption	DiD and PSM	Positive

Table 2.A.2: Literature review on technological change & health outcomes

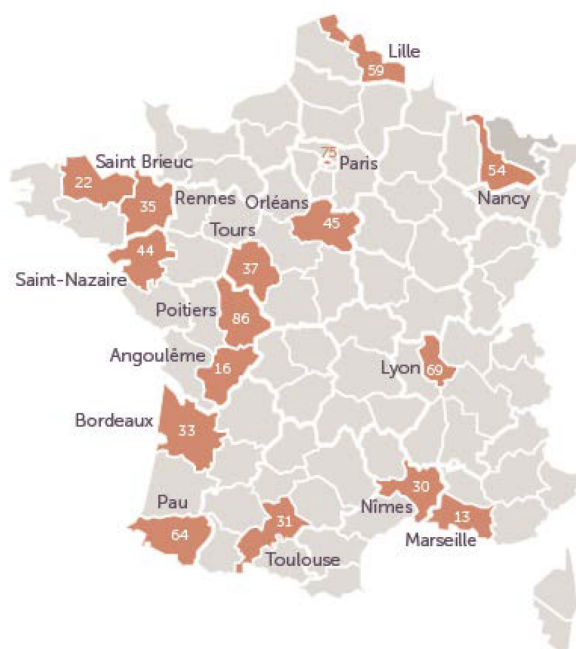
Paper	Context	Outcomes	Technology	Methods	Results on health
Arntz et al. [2024]	Germany 2011-2019	Self-reported health and Sick leaves (survey)	Computerisation & Digitalization (survey)	OLS	Negative effect on manual workers No effect on cognitive workers
Abeliansky et al. [2024]	Germany 2002-2018	Mental health (survey)	Industrial robots (IFR)	IV (robot intensity of other countries)	Negative effect
Blasco et al. [2022]	France 2013-2016	Mental health (survey)	Automation (survey)	PSM	Negative effect
Hernnäs [2023]	Sweden 1985-2015	Mortality, Morbidity (admin.)	Occupational decline	OLS	Negative effect
Gihleb et al. [2022]	USA 2005-2011 (firms, counties) Germany 1994-2016	Workplace injuries (estab.) Disability, Work accident	Industrial robots (IFR)	OLS, IV	↘ workplace injuries ↘ mental health ↘ physical job intensity, disability
Giuntella et al. [2023]	Germany 2000-2020	Well-being, Mental health (survey)	Artificial intelligence (survey)	DiD, event studies	Negative effect on job satisfaction No effect on mental health
Gunadi and Ryu [2021]	USA 2004-2017 (local labor market)	Physical health, disabilities (survey)	Industrial robots (IFR)	OLS, IV	Positive effect for low-skilled
Lordan and Stringer [2022]	Australia 2001-2018	Mental health, Life satisfaction (survey)	Automation (survey)	OLS	Negative effect

Notes: All studies are conducted at the worker level unless specified otherwise in the Context column.

2.B Data

2.B.1 Constances cohort

Figure 2.B.1: Map of targeted zones by Constances



Notes: This map plots the 16 *départements* (with partner Health Screening Centers) where the Constances population was drawn.

2.B.1.1 Definition of labor market outcomes

Administrative data on labor market outcomes come from the National Retirement Insurance Fund and provide the number of validated quarters that contribute to an individual's pension.

Quarters validation.

Table 2.B.1: Requirements to validate quarters in each status

Status	Requirements to validate 1 quarter ⁵¹
Employment	Earnings $\geq 200 \times$ gross hourly minimum wage (before 2014) Earnings $\geq 150 \times$ gross hourly minimum wage (after 2014)
Unemployment	50 days of unemployment benefits
Sick leave and work accident	60 days of compensation
Disability	90 days of compensation

⁵¹Requirements to validate quarters are presented here: <https://www.legifrance.gouv.fr/codes/id/LEGISCTA000006173397/>

- Unemployment: https://www.legislation.cnnav.fr/Pages/expose.aspx?Nom=periode_assimilee_chomage_periode_validable_ex

- Sick leave and work accident: https://www.legislation.cnnav.fr/Pages/expose.aspx?Nom=periode_assimilee_

Employment status. We define six statuses that correspond to the main employment situation for each year: "Employment", "Unemployment", "Inactivity", "Retirement", "Sick leave and work accident" and "Maternity leave". Following [Rabaté and Rochut \[2017\]](#), the computation is based on

- the number of quarters validated in each state for a given year (hierarchical order in case of a tie: employment, unemployment, sickness, maternity leave, disability, retirement and inactivity)
- with additional correction for time spent in each state when information is available (e.g. retirement date; unemployment, sickness and disability periods correspond to a minimum number of days spent in the state)

2.B.1.2 Definition of health outcomes

Administrative data on the healthcare consumption of individuals come from the Nation Health Insurance Fund.

Table 2.B.2: Definition of health outcomes

Outcome	Definition
<i>Visits</i>	<i>Visits to</i>
GP	"General practitioners", "Specialist in general medicine with diploma", "Specialist in general medicine recognized by the Order"
Mental health specialist	"Psychiatrists" and "Neuropsychiatrists"
<i>Hospital stays</i>	<i>CIM10 classification</i>
All (except pregnancy)	All except O00-O99
Cancer	C00-D48
Circulatory problems	I00-I99
Alcohol-related issues	F10, I85, K70, K860, T500, T510
Mental health	F00-F99
<i>Drugs</i>	<i>ATC7 classification</i>
Antidepressants and anxiolytics	N06A and N05B
Sleeping pills	N05C
Pain killers	N02
Opioids	N02A
Anti-inflammatories	M01

maladie_maladie_ex; https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_accident_travail_accident_travail_ex
 - Disability: https://www.legislation.cnav.fr/Pages/expose.aspx?Nom=periode_assimilee_invalidite_invalidite_ex

2.B.2 Firm-level data

2.B.2.1 Structural and financial information of companies

The FICUS and FARE datasets provide firm-level structural and financial information from administrative sources and annual surveys of companies.

The raw tax files from the French *Bénéfices Réels Normaux (BRN)* database contain detailed financial information for the universe of French firms. In particular, they provide the balance-sheet value and depreciation of "industrial equipment and machines" fixed assets in euros, which we use to construct our firm-level measure of investment in industrial capital.

In both databases, balance-sheet observations are identified by year and firm unique identifier (SIREN). We keep records of 12-month fiscal years for consistent 9-digit SIREN identifiers. Then, we apply cleaning procedures based on [Kalemli-Ozcan et al. \[2015\]](#). We drop firms that report negative sales, total assets or tangible fixed assets in any year; negative industrial assets in any year (for BRN data); negative or misreported employment above 2 million employees (for FICUS FARE data). We also drop firms that report "a value of sales to total assets larger than the 99.9 percentile of the distribution" in any year. Furthermore, we drop SIREN-year observations when sales, total assets and operating revenue are simultaneously missing; and when total assets or fixed assets are zero. In order to handle the impact of outliers, we follow the literature and winsorize all the financial ratios by industry and year to keep the kurtosis of the distribution under 10.

2.B.2.2 Identification of automation goods in Customs

Table 2.B.3: HS-2012 product codes referring to automation technologies and robots [Acemoglu and Restrepo, 2022]

Label	HS-2012 codes
1. Industrial robots	847950
2. Dedicated machinery	847989
3. Automatic machine tools (incl. Numerically controlled machines)	845600–846699, 846820–846899, 851511–851519
4. Automatic welding machines	851521, 851531, 851580, 851590
5. Weaving and knitting machines	844600–844699, 844700–844799
6. Other textile dedicated machinery	844400–844590
7. Automatic conveyors	842831–842839
8. Automatic regulating instruments	903200–903299

Notes: This table comes directly from Domini et al. [2022].

Harmonization. Given that the Combined Nomenclature has been regularly revised during the period of interest, an algorithmic procedure is implemented to harmonize product codes over time and allow for panel data analysis at a disaggregated product level. We use the algorithm from Behrens and Martin [2015] which consists of aggregating NC8-year codes into new time-invariant "products" that contain the smallest set of CN8-year codes linked together through revisions of the nomenclature during the period. See Bergounhon et al. [2018] for a description of the procedure. A product is considered as an *automation product* if it contains any of the codes from Acemoglu and Restrepo [2022] for the year 2012. This ensures that variation in the flows of products does not merely reflect changes in the definition of CN8 products over time, which would be spuriously interpreted as fluctuations in firm-level trade of products. Mechanically, the longer the time period of harmonization, the more CN8-year products are aggregated in the new "product" definition, especially when the considered period includes large revisions of the nomenclature. This means that there might be a trade-off between identifying automation products more precisely and extending the analysis over a longer time span. For the sake of our analysis, we harmonize the product classification over the period 2002-2019; this allows us to observe 5 years of imports prior to our period of interest to construct our proxies, while avoiding lumping even more codes under the automation products due to the 2002 HS revision.

2.B.2.3 Proxies for the stock of imported automation goods

In order to study variations in the net stock of automation goods in the firm, we compute two additional proxies that are based on past imports.⁵² Using a perpetual inventory method, the net stock of automation technologies of firm j in year t (Stock_{jt}) is computed in the following way

- with a proportional depreciation rate, inspired by [Graetz and Michaels \[2018\]](#) and [Bonfiglioli et al. \[2024\]](#)

$$\text{Stock}_{jt} = \sum_{\substack{k \leq t \\ k \geq 2002}} (1 - 0.15)^{(t-k)} \times AI_{jk} \quad (2.B.1)$$

- with a linear depreciation rate over 8 years⁵³, inspired by [Aghion et al. \[2024\]](#)

$$\text{Stock}_{jt} = \sum_{\substack{k=t-7 \\ k \geq 2002}}^t \left(1 - \frac{(t-k)}{8}\right) \times AI_{jk} \quad (2.B.2)$$

where AI_{jk} is the flow of imports of automation technologies of firm j in year t .

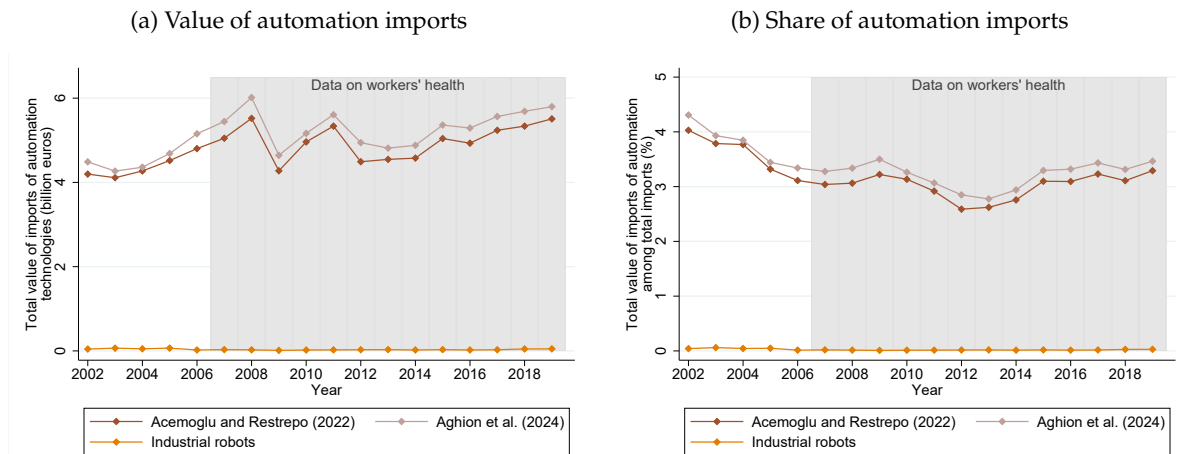
⁵²Note that the distribution of the log change in the stock is more closely aligned with the stock reported in balance-sheet data when linear depreciation is applied.

⁵³The depreciation is over 8 years for machinery and equipment, following French accounting standards for amortization. Given that our treatment window starts in 2009, our measure will incorporate information on imports from 2002 to 2019.

2.C Additional graphs for the stylized facts

2.C.1 Fact 1

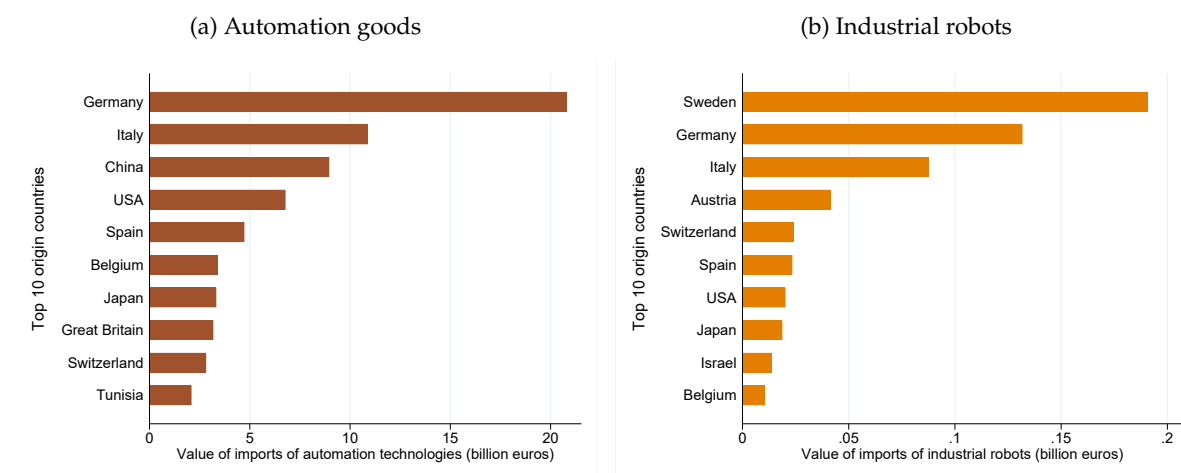
Figure 2.C.1: Macro trends in automation imports



Notes: These figures plot the evolution of imports of automation goods between 2002 and 2019. Computations are made on the sample of eligible firms. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The light brown line corresponds to imports of products from the list in [Aghion et al. \[2024\]](#), the brown line to the list in [Acemoglu and Restrepo \[2022\]](#), and the orange line to industrial robots only. Outcomes: (a) Total value of imports of automation technologies (billion 2010 euros); (b) Share of imports of automation technologies among total imports (%).

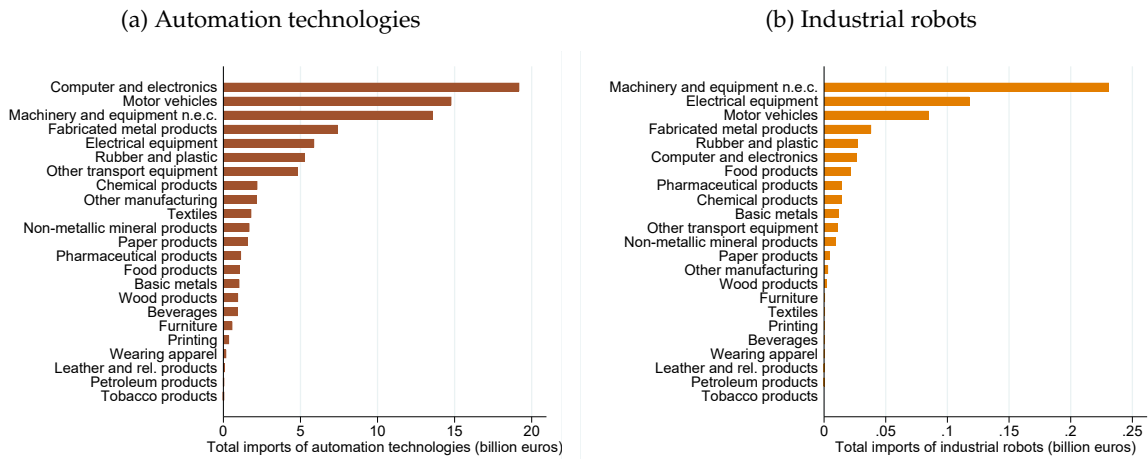
2.C.2 Fact 2

Figure 2.C.2: Main origin countries of automation technologies and industrial robots



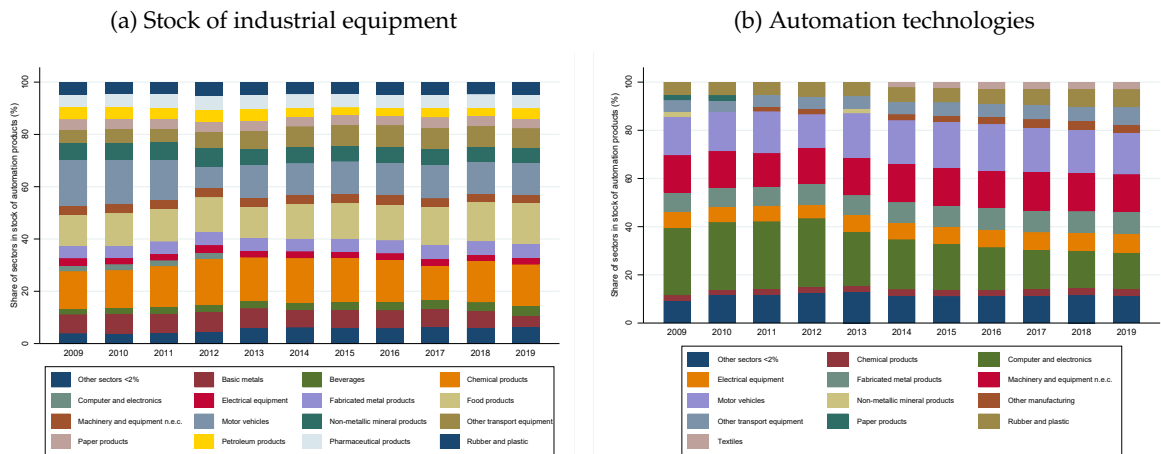
Notes: These graphs show the total imports of (a) automation technologies ([Acemoglu and Restrepo \[2022\]](#)'s measure) and (b) industrial robots by origin countries. Computations are made on the sample of eligible firms, computed from 2002 to 2019.

Figure 2.C.3: Imports of automation technologies by French industry



Notes: These graphs show the total imports of (a) automation technologies (Acemoglu and Restrepo [2022]’s measure) and (b) industrial robots by industries. Computations are made on the sample of eligible firms, computed from 2002 to 2019.

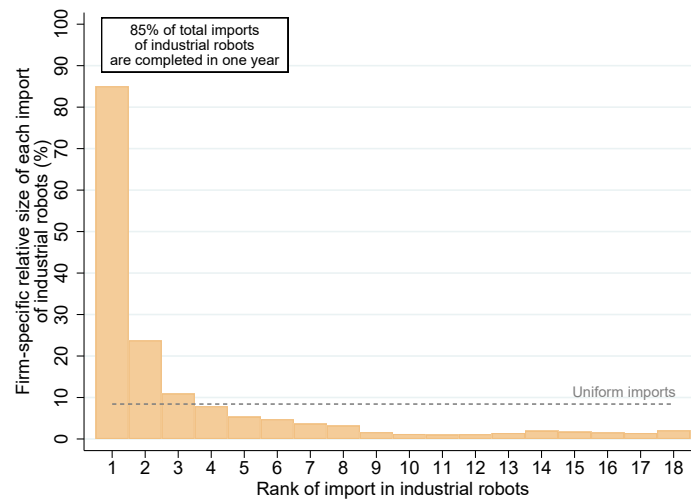
Figure 2.C.4: Stock of automation capital by sector



Notes: These graphs plot the share of sectors in the stock of (a) industrial equipment and machines (proxy 1) and (b) automation goods (Acemoglu and Restrepo [2022]’s measure) in the sample of eligible firms from 2009 to 2019. The proxy stock of automation goods (Acemoglu and Restrepo [2022]’s measure) is computed from the flows of imports using a linear depreciation over 8 years.

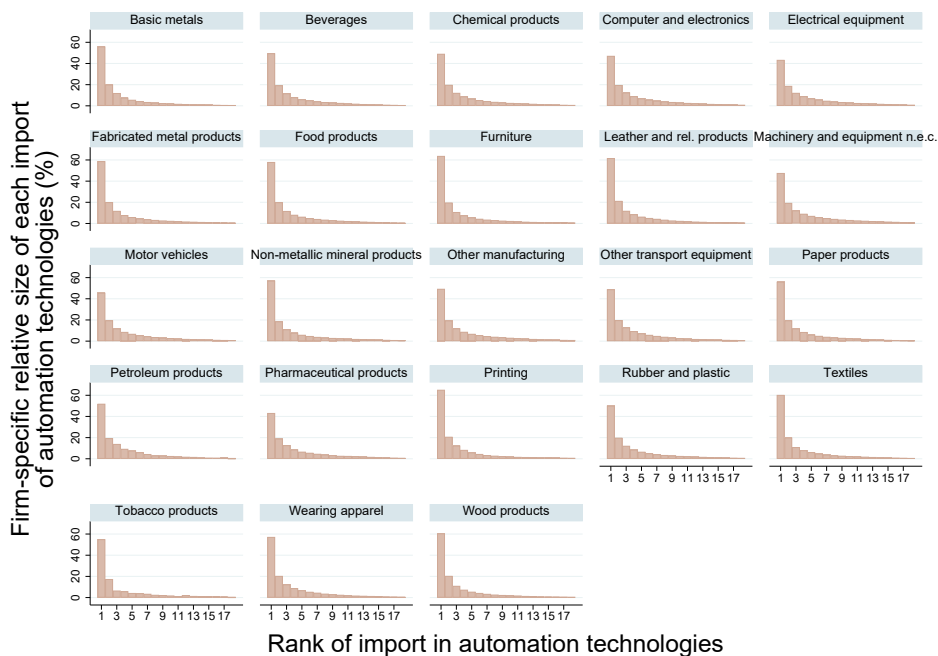
2.C.3 Fact 3

Figure 2.C.5: Relative size of the import according to its rank - industrial robots



Notes: This graph illustrates the relative size of each import according to its rank. The y-axis displays the share of each import in the firm's total imports over the period; considering only imports of industrial robots. The x-axis represents the rank (according to its value in 2010 euros) of each import. The dotted line represents the hypothetical scenario in which firms would import uniformly, *i.e.*, if imports were perfectly smoothed over time, throughout the entire period. Computations are based on a sample of eligible firms, from 2002 to 2019. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment.

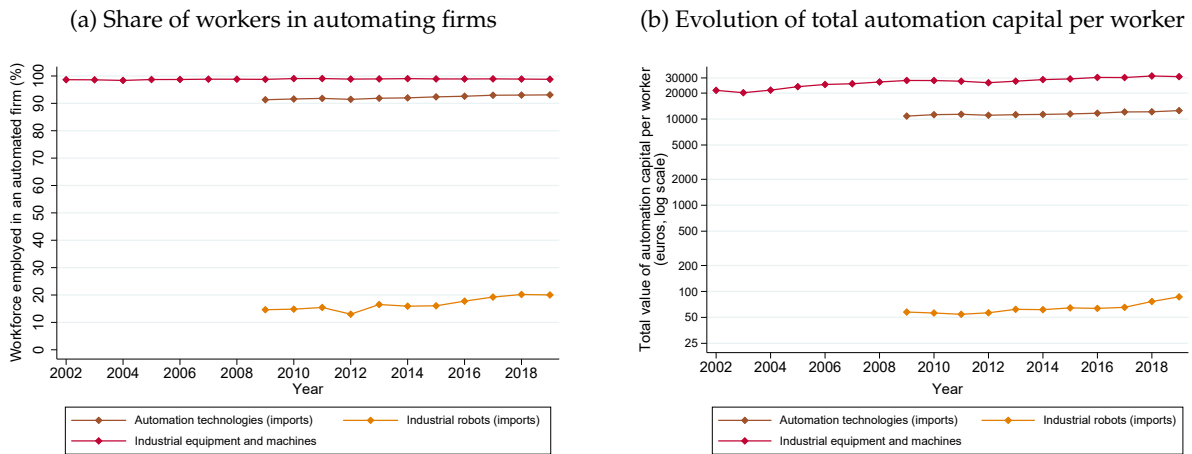
Figure 2.C.6: Relative size of the import according to its rank - by sector



Notes: This graph illustrates the relative size of each import according to its rank across sectors. The y-axis displays the share of each import in the firm's total imports over the period; considering only imports of automation products. The x-axis represents the rank (according to its value in 2010 euros) of each import. Computations are based on a sample of eligible firms, from 2002 to 2019. Values are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment.

2.C.4 Fact 4

Figure 2.C.7: Workers' exposure to automation



Notes: These graphs display the exposition of the workforce to firm-level automation on the sample of eligible firms. (a) This graph displays the share of the workforce that is employed in a firm with a positive stock of automation capital. (b) This graph plots the ratio of the stock of automation capital to the total workforce for our three proxies. The measures of net stock of automation capital are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The proxy stocks of automation technologies (Acemoglu and Restrepo [2022]'s measure) and industrial robots are computed from the flows of imports using a linear depreciation over 8 years.

2.C.5 Stylized facts for the narrowed sample of firms

In the narrow sample of firms, the aggregate net balance-sheet value of tangible fixed assets has also increased between 2004 and 2008 but has been overall more stable afterward. Meanwhile, the total value of the automation stock has increased continuously over the period and therefore, so has the ratio of automation to total industrial equipment stock. Of course, this partly reflects the fact that this sample is restricted to continuously operating firms (stabilizing the total value of assets around the crisis) and that we select them only if they have an import of automation over the period.

The stock of automation capital has also remained stable in the sample of eligible firms, even though it has increased in the narrow sample of continuously operating firms that automate. These aggregate trends do not show an increase in the volume of machines per worker, but the narrow sample selects firms with a rising share of automation capital on average.

While the total value of industrial equipment per worker is similar in the sample of eligible firms and the narrow sample (around 30,000 euros₂₀₁₀), we observe that the intensity of automation capital per worker becomes slightly higher after an increase in

the narrow sample (from around 11,000 euros₂₀₁₀ in 2009 to about 14,000 euros₂₀₁₀ in 2016).

2.D Definition of exposure to automation

2.D.1 Discussion on the definition of automation event

Various shocks have been exploited in the literature.

First investment. An obvious shock to consider is the first adoption of industrial machines [Bonfiglioli et al., 2024].

Limitations. This shock works well when such events are relatively rare. Indeed, we want to avoid scenarios where a firm adopts a technology merely to test it, but only truly undergoes a technological shift later when it invests heavily in that technology. Additionally, from import data, we cannot be certain of capturing the initial investment (which might have occurred earlier or been purchased domestically in France).

Top investment/Spike. Another strategy is to consider the biggest investment [Domini et al., 2022] or a similar but more restrictive measure imposing a size threshold [Bessen et al., 2023]. We adapt the spike definition in Bessen et al. [2023] (computed on automation costs): a firm j is subject to a spike in the imports of automation goods in year t if the automation import ratio in year t (defined as the value of automation imports, relative to average annual imports excluding automation goods) is at least three times higher than the average automation import ratio observed in the other years.⁵⁴

Formally,

$$\text{Spike}_{jt} = \mathbb{1} \left\{ \frac{AI_{jt}}{\overline{TI}_j} \geq 3 \times \frac{1}{T-1} \sum_{k \neq t} \frac{AI_{jk}}{\overline{TI}_j} \right\}$$

where AI_{jt} is the value of imports of automation technologies of firm j in year t and \overline{TI}_j is the average total annual imports (excluding automation goods) of firm j across all years.

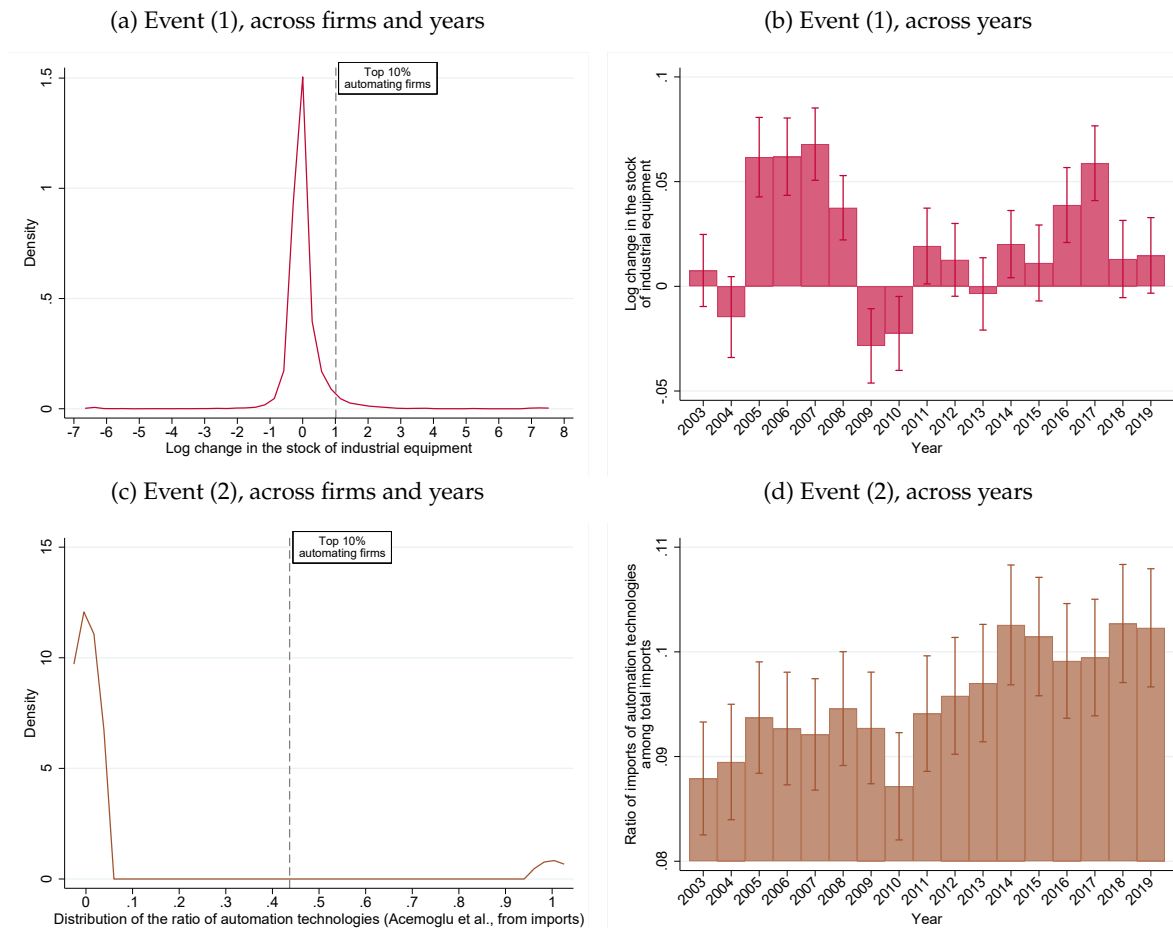
Limitations. To characterize this type of event, it is necessary to account for all periods of observation. This means that the event's definition is based on what will happen

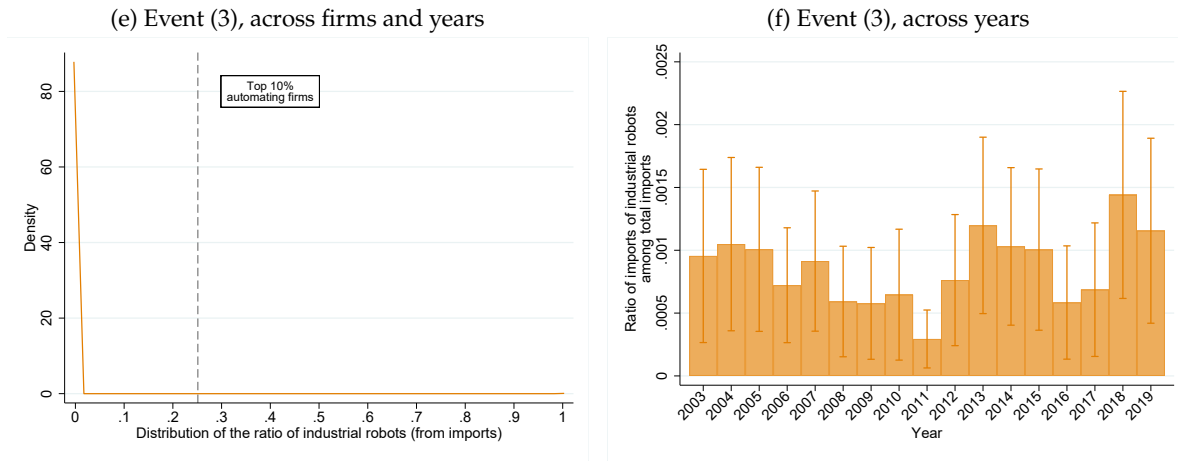
⁵⁴In the case of firms importing only robots throughout the entire period, the denominator \overline{TI}_j being equal to zero, we define the spike year as the year with the highest import of automation technologies.

afterward. We might pinpoint the moment when the firm reaches its peak productivity, especially because it will not make a larger investment later. This poses a particular issue because any other outcomes (such as sales or employment) are likely to also show a mechanical decline after the event. Additionally, with the current sample definition, it is not possible to use a set of workers who were never treated (due to the restriction that firms must have imported automation technology over the period). We address this limitation by relaxing the assumption when examining these two specific shocks.

2.D.2 Additional graphs on the definition of treatment

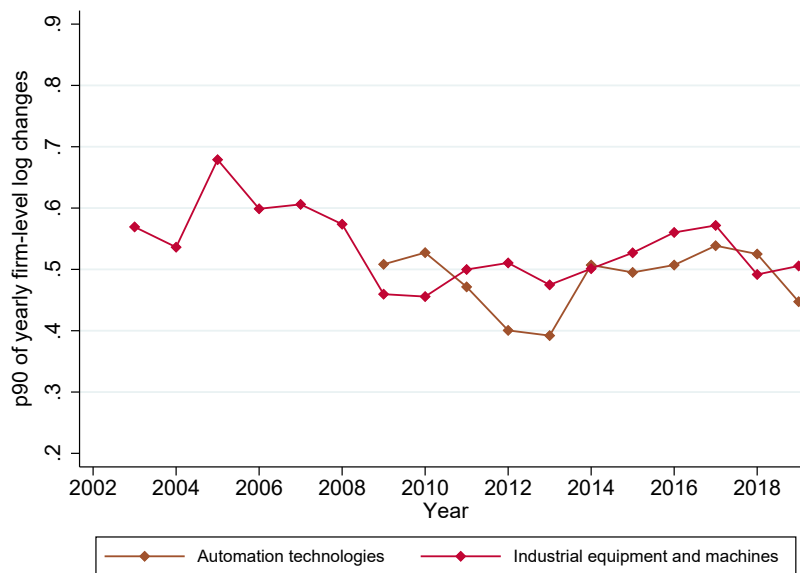
Figure 2.D.1: Distribution of the log changes





Notes: The figures plot the distribution of the different investment events across firms and years and across years only. The first panel is for the investment defined as a log change in the stock of industrial equipment and machines. The second panel is for the investment defined as the ratio of imports of automation goods (Acemoglu and Restrepo [2022]’s definition) over all imports. The third panel is for the investment defined as the ratio of industrial robots over all imports.

Figure 2.D.2: Evolution of the p90 of log changes in automation stock over time



Notes: This graph plots the yearly 90th percentile of the distribution of log changes in the annual stock of automation capital. The measures of net stock of automation capital are converted to constant euros from 2010 using the deflator of gross fixed capital formation of machinery and equipment. The 90th percentile stock of industrial robots cannot be included as there are too few positive observations per year.

2.D.3 Sample size according to the treatment definition

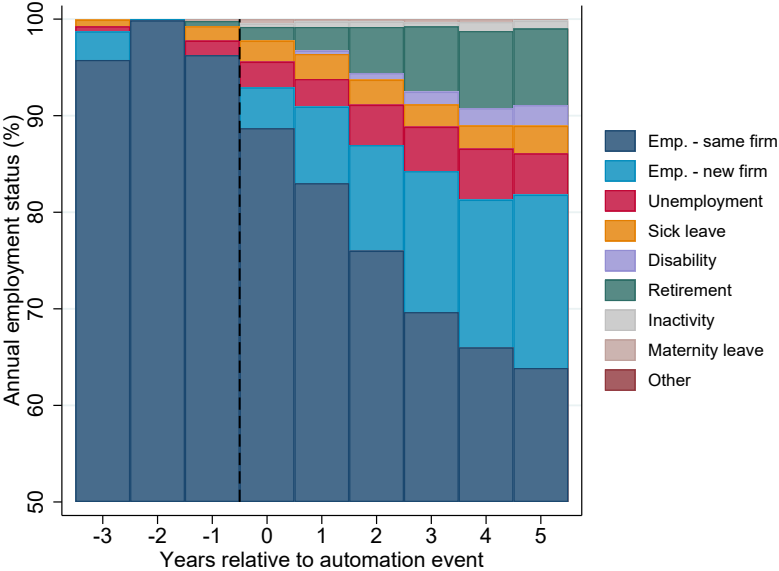
Table 2.D.1: Sample composition

		Firms	
		Treated	Control
Investment: log change in the stock			
(1)	Industrial equipment (proxy 1)	1,518	3,784
	Industrial equipment (proxy 1), rescaled by size	1,296	4,087
	Automation tech. (proxy 2a)	1,395	4,031
	Automation tech. (proxy 2a, lin.)	1,351	3,982
	Industrial robots (proxy 2b)	29	5,728
	Industrial robots (proxy 2b, lin.)	29	5,728
Investment: log change in the intensity			
	Automation tech. capital intensity (stock of proxy 2a / tangible fixed assets)	1,185	4,342
	Automation tech. capital intensity (stock of proxy 2a, lin. / tangible fixed assets)	1,029	4,516
	Automation tech. equipment intensity (stock of proxy 2a / industrial equipment)	1,369	4,128
	Automation tech. equipment intensity (stock of proxy 2a, lin. / industrial equipment)	1,190	4,306
	Equipment capital intensity	1,515	3,869
Investment: flows of imports			
(2)	Automation tech.	1,164	4,352
	Automation tech. (over net tangible fixed assets)	801	4,835
(3)	Industrial robots	16	5,747
Investment: first import			
(3)	Industrial robots	188	5,500

Notes: This table displays the size of the treated group (column 2) and control group (column 3) for each treatment definition on the narrow sample of firms used for regressions.

2.D.4 Additional descriptive statistics

Figure 2.D.3: Employment status and time to event



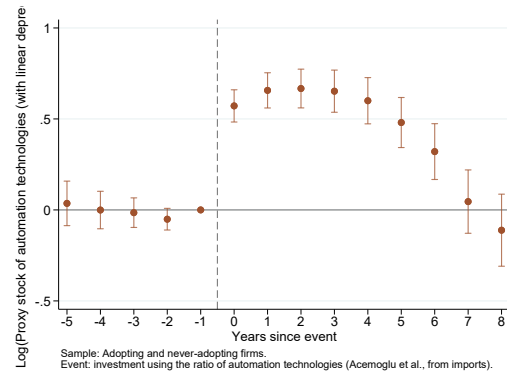
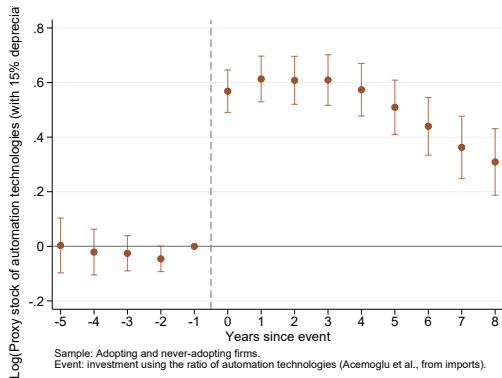
Notes: This Figure plots the employment composition of the treated group in each year relative to the automation event (year 0). The automation event is defined based on large investments in industrial equipment and machines (proxy). The categories "Employment - Same firm" and "Employment - New firm" refer to changes in the main employer compared to year -1. The procedure to define the main employment status is described in Appendix 2.B.1.1.

2.E Additional graphs for the Firm-level results

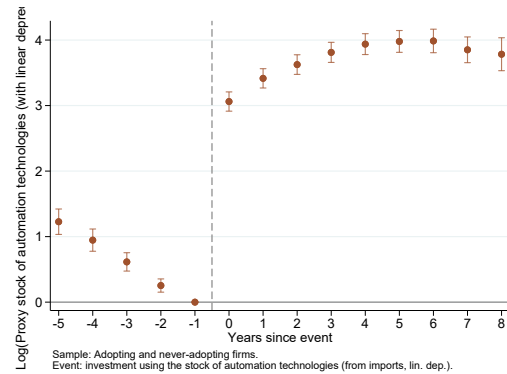
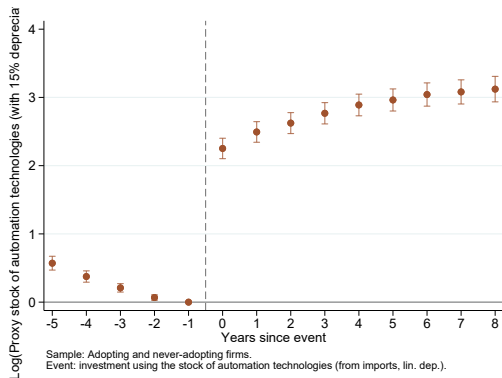
2.E.1 Validation exercise

Figure 2.E.1: Event studies at the firm level - results on automation

(a) Proxy stock of automation tech. (imports, 15% de- (b) Proxy stock of automation tech. (imports, linear
preciation). Investment: ratio (2) depreciation). Investment: ratio (3)

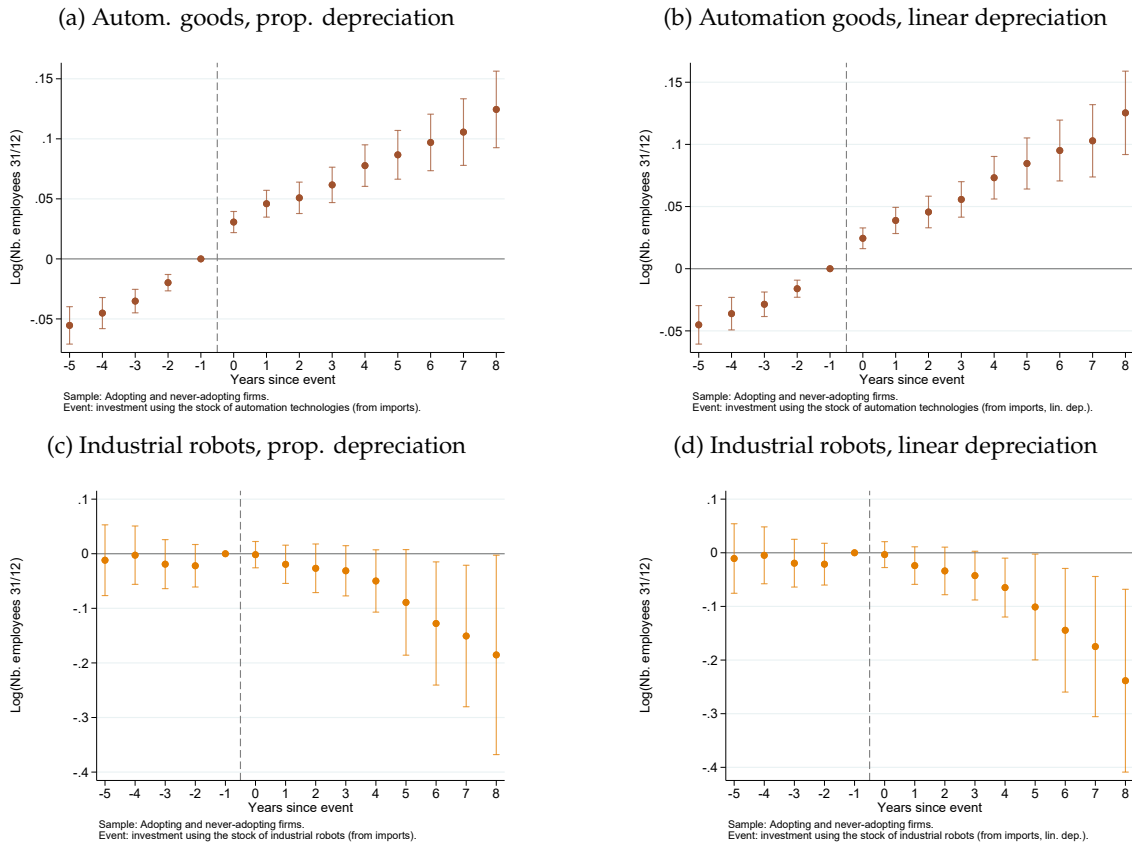


(c) Proxy stock of automation tech. (imports, 15% de- (d) Proxy stock of automation tech. (imports, linear
preciation). Investment: stock variation dep.). Investment: stock variation



Notes: These figures plot the association of automation with different outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment: (a) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]'s measure, proxy stock from imports with 15% depreciation); (b) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]'s measure, proxy stock from imports with linear depreciation); (c) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]'s measure, proxy stock from imports with 15% depreciation); (d) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]'s measure, proxy stock from imports with linear depreciation). Outcome: (a) Log(proxy stock of automation technologies, from Acemoglu and Restrepo [2022]'s measure, with 15% depreciation); (b) Log(proxy stock of automation technologies, from Acemoglu and Restrepo [2022]'s measure, with linear depreciation); (c) Log(proxy stock of automation technologies, from Acemoglu and Restrepo [2022]'s measure, with 15% depreciation); (d) Log(proxy stock of automation technologies, from Acemoglu and Restrepo [2022]'s measure, with linear depreciation). The standard errors are clustered at the firm level.

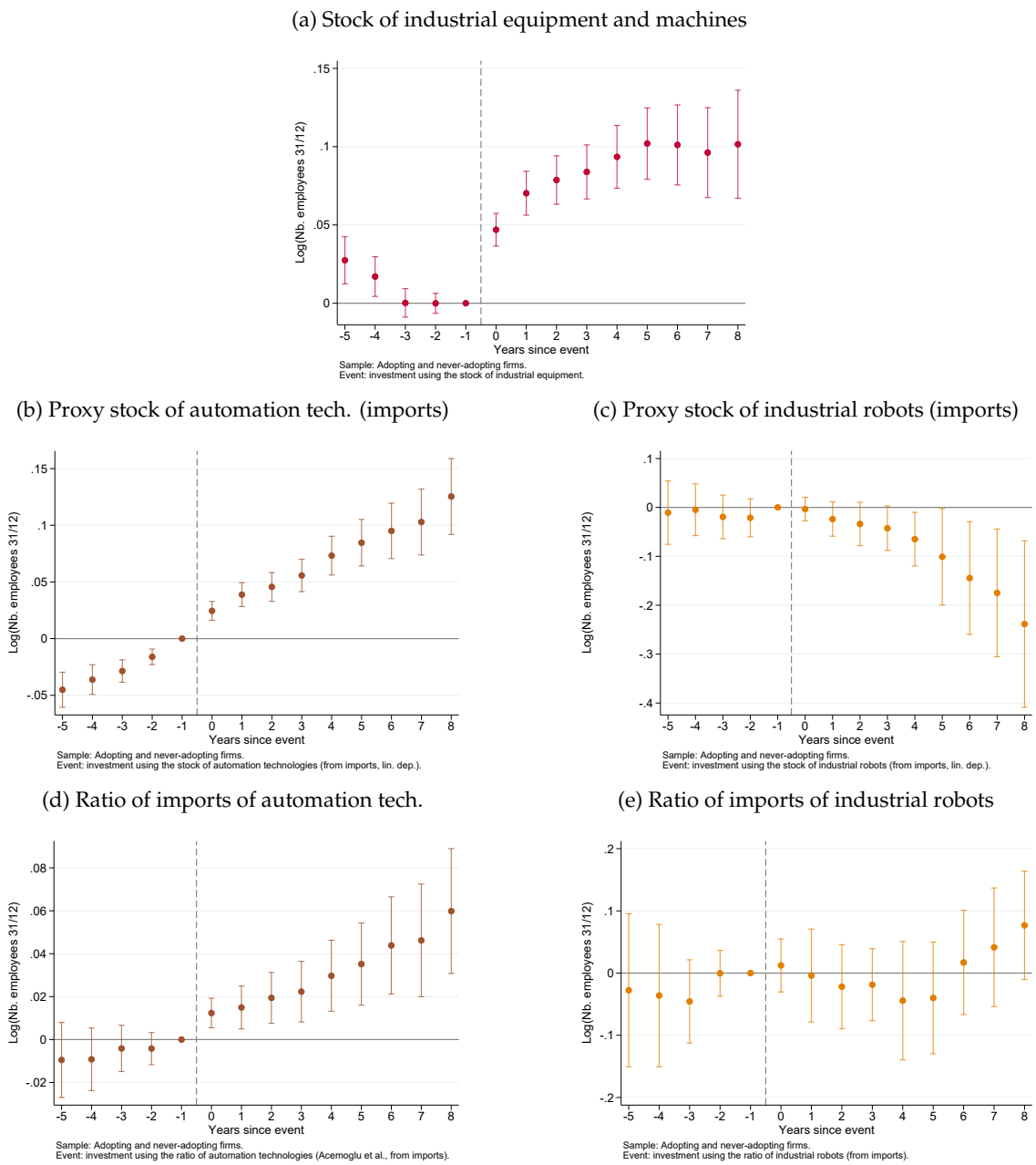
Figure 2.E.2: Event studies at the firm level (results on employment) - comparison of different proxies for the stock of goods



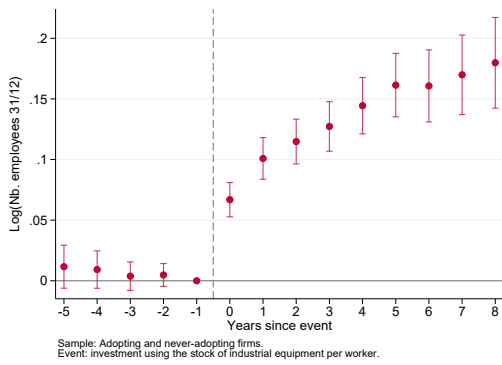
Notes: These figures plot the association of different automation events with different proxies for the stock of goods. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment is measured as the log change in: (a) the proxy stock of automation technologies (Acemoglu and Restrepo [2022]' measure), with proportional depreciation from equation 2.B.1; (b) the proxy stock of automation technologies (Acemoglu and Restrepo [2022]' measure), with linear depreciation from equation 2.B.2; (c) the proxy stock of industrial robots, with proportional depreciation from equation 2.B.1; (d) the proxy stock of industrial robots, with linear depreciation from equation 2.B.2. The standard errors are clustered at the firm level.

2.E.2 Changing automation event

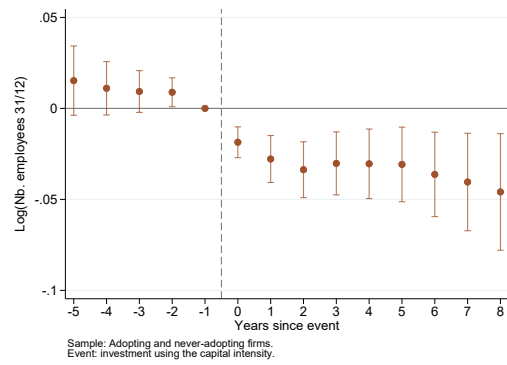
Figure 2.E.3: Event studies at the firm level (results on employment) - comparison of different measures



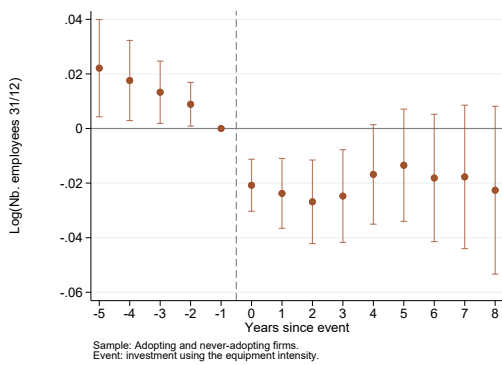
(f) Intensity: stock industrial equipment / size



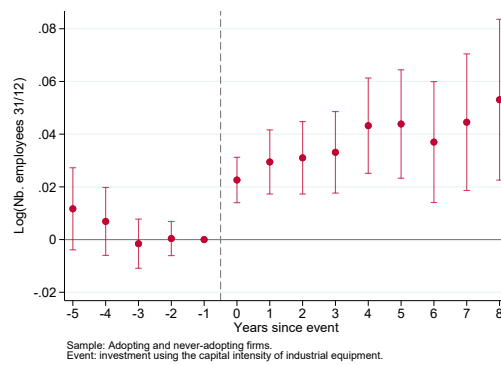
(g) Intensity: stock autom. tech. / stock physical capital



(h) Intensity: stock autom. tech. / stock industrial equipment



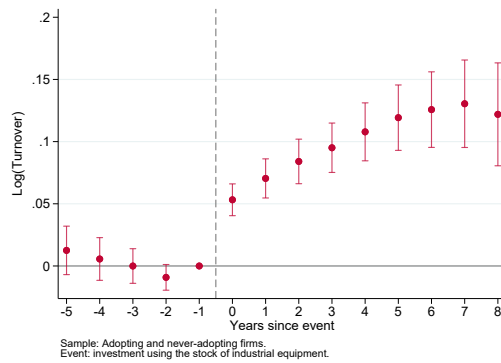
(i) Intensity: stock industrial equipment / stock physical capital



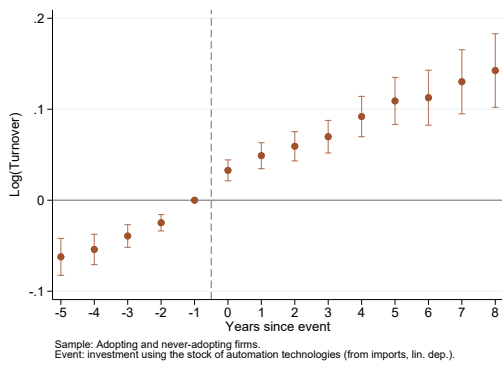
Notes: These figures plot the association of different automation events with the size of firms. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment measured as: (a) Log change in the stock of industrial equipment and machines; (b) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports); (c) Log change in the stock of industrial robots (proxy stock from imports); (d) Ratio of imports of automation technologies over total imports (Acemoglu and Restrepo [2022]' measure); (e) Ratio of imports of industrial robots over total imports; (f) Intensity as variation in stock of industrial equipment over lagged size of the firm; (g) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of tangible fixed assets; (f) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of industrial equipment and machines; (i) Intensity as stock of industrial equipment and machines over stock of tangible fixed assets (from BRN). The standard errors are clustered at the firm level.

Figure 2.E.4: Event studies at the firm level (results on sales) - comparison of different measures

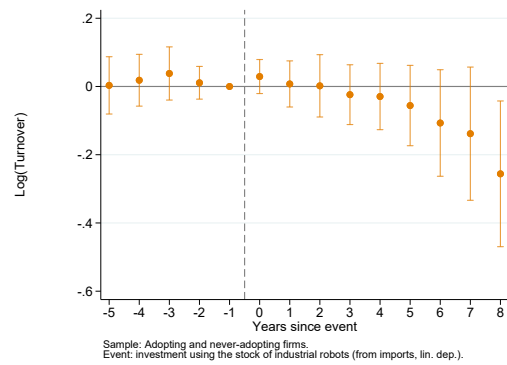
(a) Stock of industrial equipment and machines



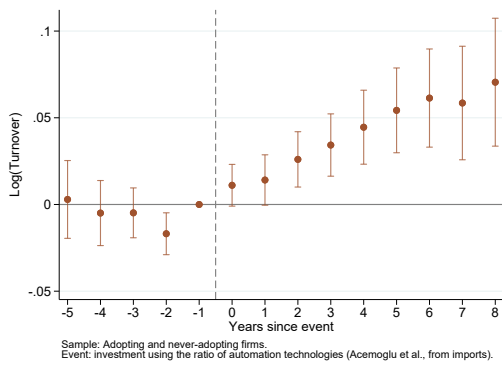
(b) Proxy stock of automation tech. (imports)



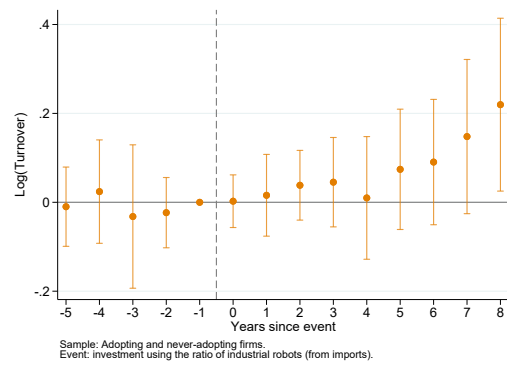
(c) Proxy stock of industrial robots (imports)



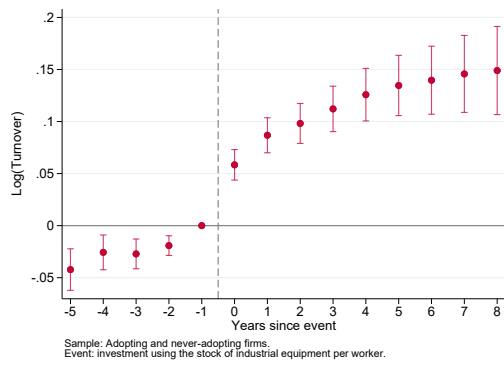
(d) Ratio of imports of automation tech.



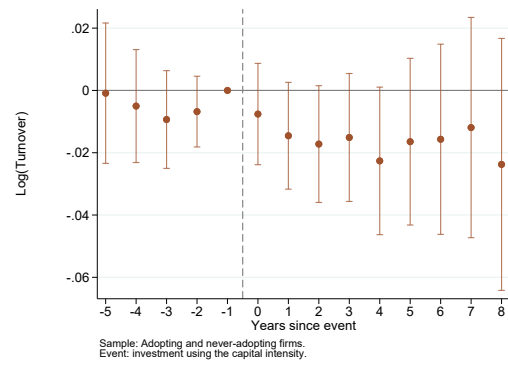
(e) Ratio of imports of industrial robots



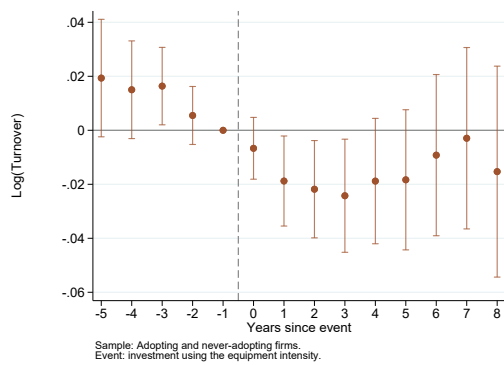
(f) Intensity: stock industrial equipment / size



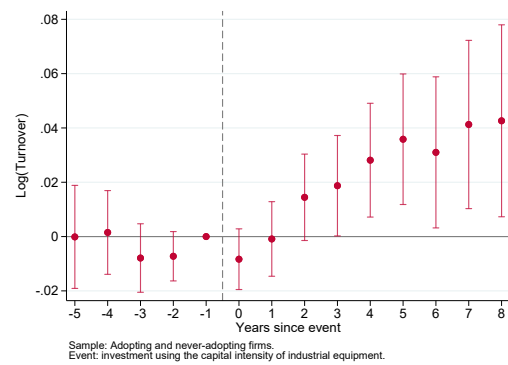
(g) Intensity: stock autom. tech. / stock physical capital



(h) Intensity: stock autom. tech. / stock industrial equipment



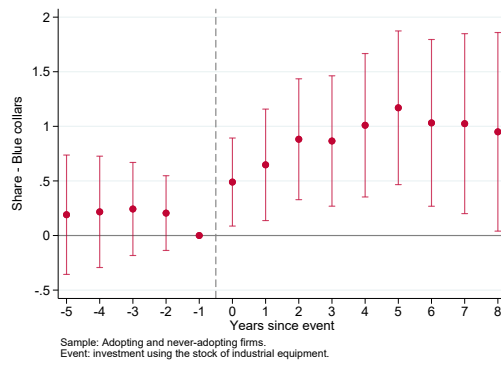
(i) Intensity: stock industrial equipment / stock physical capital



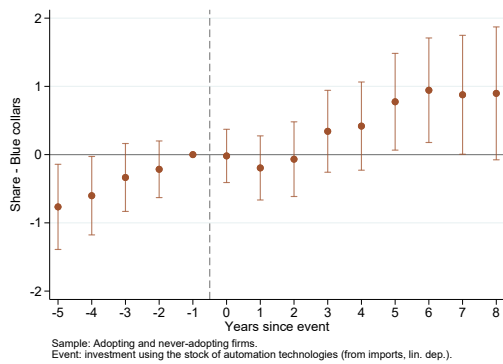
Notes: These figures plot the association of different automation events with the sales of firms. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment measured as: (a) Log change in the stock of industrial equipment and machines; (b) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports); (c) Log change in the stock of industrial robots (proxy stock from imports); (d) Ratio of imports of automation technologies over total imports (Acemoglu and Restrepo [2022]' measure); (e) Ratio of imports of industrial robots over total imports; (f) Intensity as variation in stock of industrial equipment over lagged size of the firm; (g) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of tangible fixed assets; (f) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of industrial equipment and machines; (i) Intensity as stock of industrial equipment and machines over stock of tangible fixed assets (from BRN). The standard errors are clustered at the firm level.

Figure 2.E.5: Event studies at the firm level (results on the share of blue-collar workers) - comparison of different measures

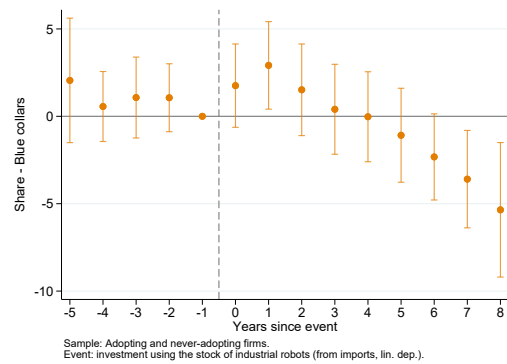
(a) Stock of industrial equipment and machines



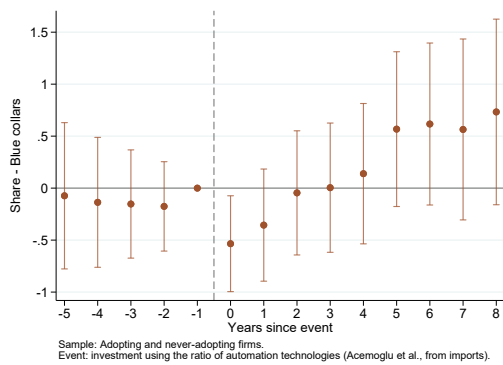
(b) Proxy stock of automation tech. (imports)



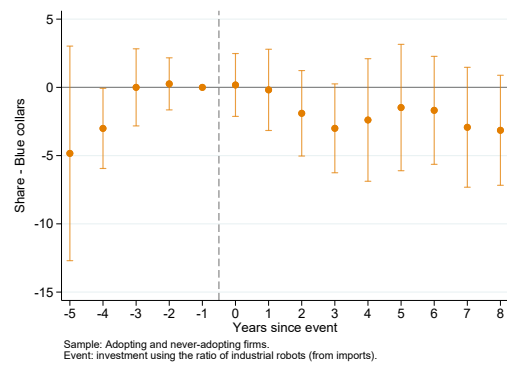
(c) Proxy stock of industrial robots (imports)



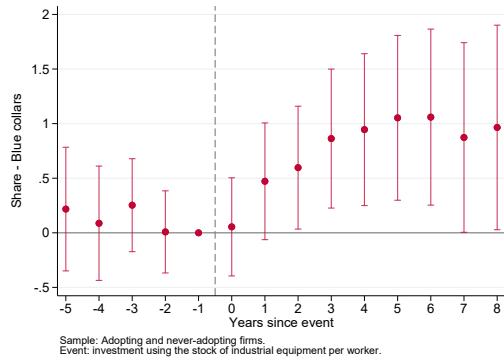
(d) Ratio of imports of automation tech.



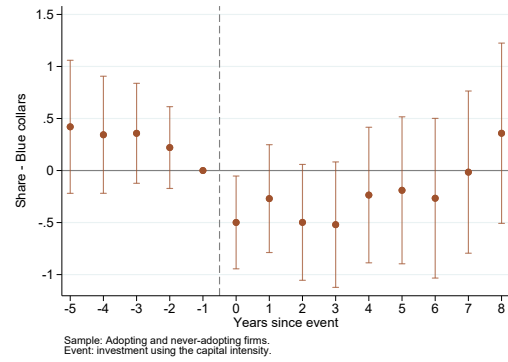
(e) Ratio of imports of industrial robots



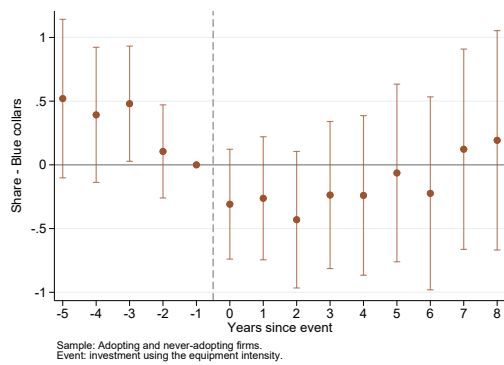
(f) Intensity: stock industrial equipment / size



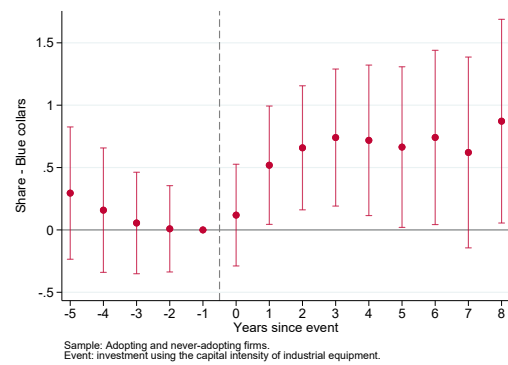
(g) Intensity: stock autom. tech. / stock physical capital



(h) Intensity: stock autom. tech. / stock industrial equipment

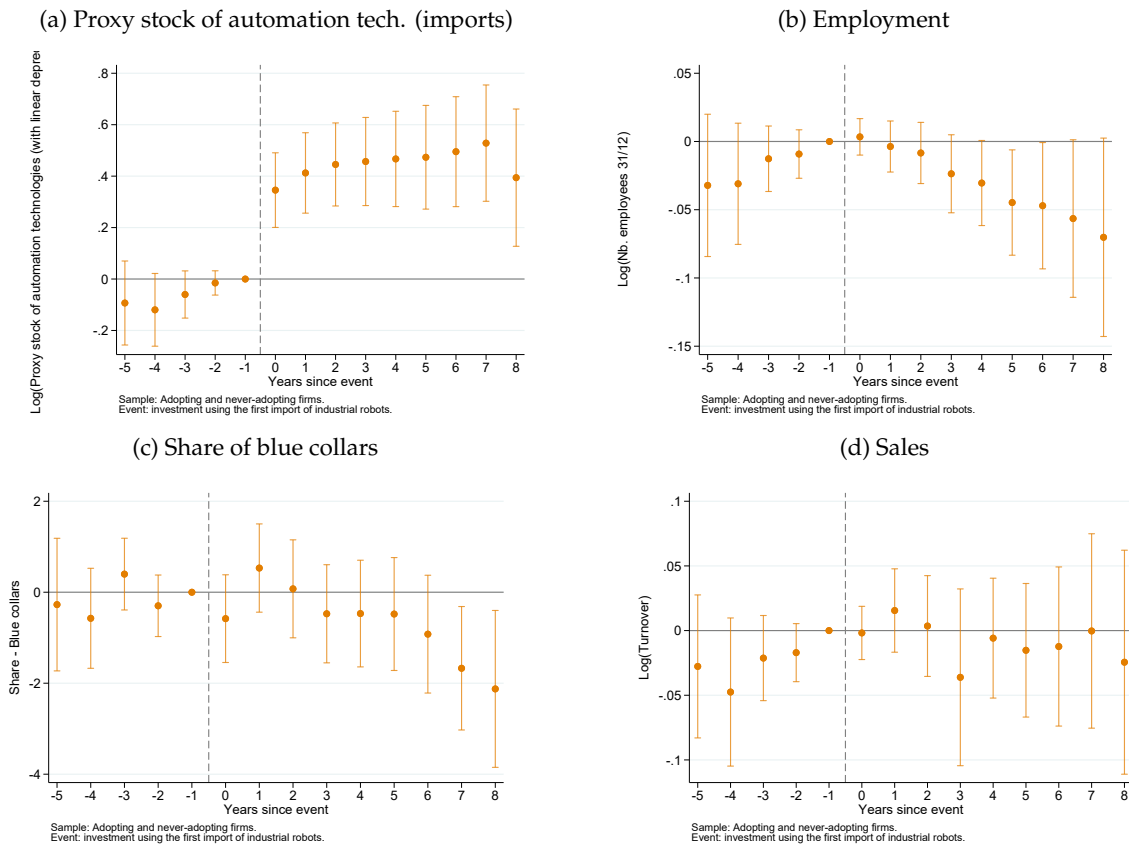


(i) Intensity: stock industrial equipment / stock physical capital



Notes: These figures plot the association of different automation events with the share of blue-collar workers. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment measured as: (a) Log change in the stock of industrial equipment and machines; (b) Log change in the stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports); (c) Log change in the stock of industrial robots (proxy stock from imports); (d) Ratio of imports of automation technologies over total imports (Acemoglu and Restrepo [2022]' measure); (e) Ratio of imports of industrial robots over total imports; (f) Intensity as variation in stock of industrial equipment over lagged size of the firm; (g) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of tangible fixed assets; (f) Intensity as stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports) over stock of industrial equipment and machines; (i) Intensity as stock of industrial equipment and machines over stock of tangible fixed assets (from BRN). The standard errors are clustered at the firm level.

Figure 2.E.6: Event studies at the firm level - treatment: first import of industrial robots



Notes: These figures plot the association of the automation event defined as the first import of industrial robots with firms' outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions of equation 2.1. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Investment is measured as the first import of industrial robots. Outcomes (a) Stock of automation technologies (Acemoglu and Restrepo [2022]' measure, proxy stock from imports); (b) log(Number of employees); (c) Share of blue collar-workers; (d) Log(sales). The standard errors are clustered at the firm level.

2.E.3 Changing specification

We also consider a specification where we keep only ever-treated firms, which are more likely to be similar. We run event studies on treated and later-treated firms in a stacked difference-in-difference design.⁵⁵ We employ a stacked difference-in-differences methodology, in the spirit of Cengiz et al. [2019]. In this literature Bessen et al. [2023] rely on a similar design. To set up the panel, we create separate datasets for each cohort of firms based on their automation event year c with $c \in \{2009, \dots, 2014\}$. We denote τ the

⁵⁵Another alternative would be to use only ever-treated firms in a staggered difference-in-difference design.

$$\log(Y_{jt}) = \alpha_j + \lambda_{jt} + \sum_{\substack{k=-5 \\ k \neq \{-1, -5\}}}^8 \delta_k D_{jt}^k + \epsilon_{jt} \quad (2.E.3)$$

However, recent developments in the difference-in-differences literature have highlighted some limitations of this methodology, especially when treatment effects are heterogeneous across groups and over time [de Chaisemartin and D'Haultfoeuille, 2020a; Goodman-Bacon, 2021]. The issue arises from the forbidden comparisons where early treated firms are used as control units.

event time, *i.e.*, the calendar year t minus the automation event year c . We also include the years around the automation event $\tau \in \{-3, \dots, 4\}$ (event window). Then for each dataset, we add the same calendar years of firms that have their automation event in year $c + 5$ or later as control group.⁵⁶ For instance, the first cohort includes firms that automated for the first time in 2009. The event window for these firms spans from 2005 to 2013. The control group for this cohort comprises firms observed within the same event window (2009-2013) but that started automating only after 2014, ensuring they remain "clean controls" without any automation event during the window. This process is repeated for each cohort, and the resulting cohort-specific datasets are then stacked to align based on event time.

We run the following specification:

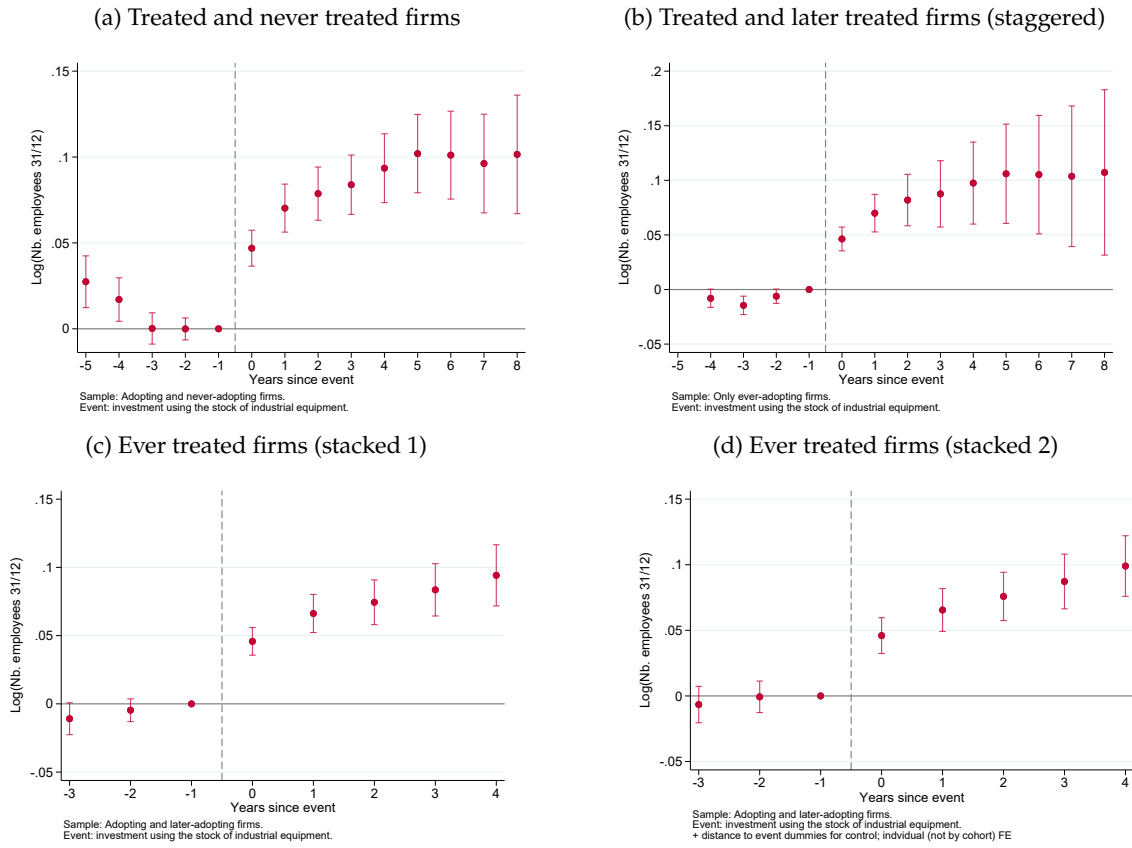
$$\log(Y_{jt}) = \alpha_j + \lambda_{st} + \sum_{\substack{k=-5 \\ k \neq -1}}^8 \gamma_k D_{jt}^k + \sum_{\substack{k=-5 \\ k \neq -1}}^8 \delta_k D_{jt}^k \cdot D_j + \xi_{jt} \quad (2.E.4)$$

where j indices firm-by-cohort, Y_{jt} is the outcome of interest of firm j in year t . D_{jt}^k are distance-to-event dummies for firm j in year t (equal to 1 if year t is k years apart from the event year, and 0 otherwise), formally $D_{jt}^k = \mathbb{1}\{t = E_j + k\}$ with E_j the event year of firm j . α_j are firm fixed effects, controlling for firms' constant characteristics over the period. λ_{st} are 5-digit industry-by-year fixed effects, controlling for industry-specific time-varying shocks. We also include γ_k coefficients associated to time-to-event dummies for the control group in order to control for specific relative trends around the event (γ_{-1} are normalized to 0). ξ_{jt} is an error term. Standard errors are clustered at the firm level. The parameters of interest are the δ_k 's, which capture the association between investment in automation technologies and firm-level outcomes, *i.e.*, the evolution of the outcomes of treated firms k years after the event compared to the year before the event (δ_{-1} is normalized to zero), relative to the evolution of control firms.

where

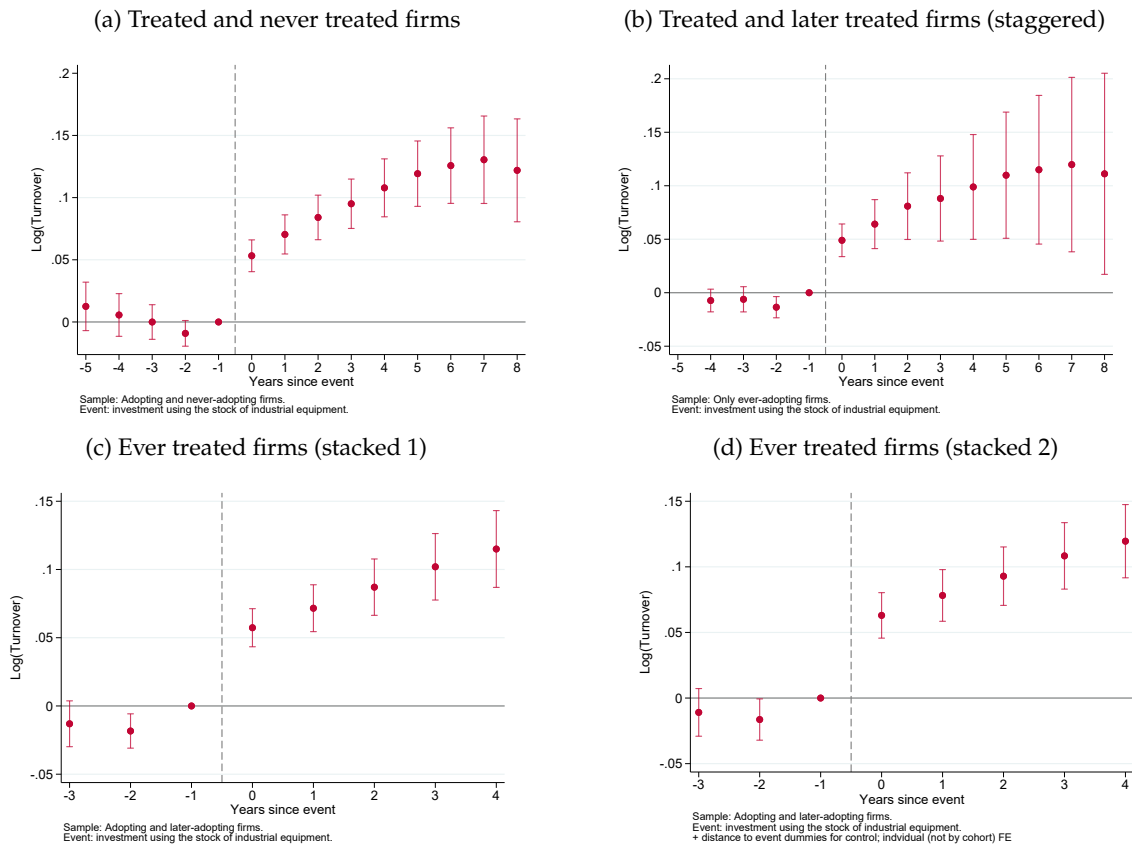
⁵⁶The treated group will be made of firms automating between 2009 and 2014 and the control group of firms automating between 2015 and 2019.

Figure 2.E.7: Event studies at the firm level (results on employment) - comparison of different specifications



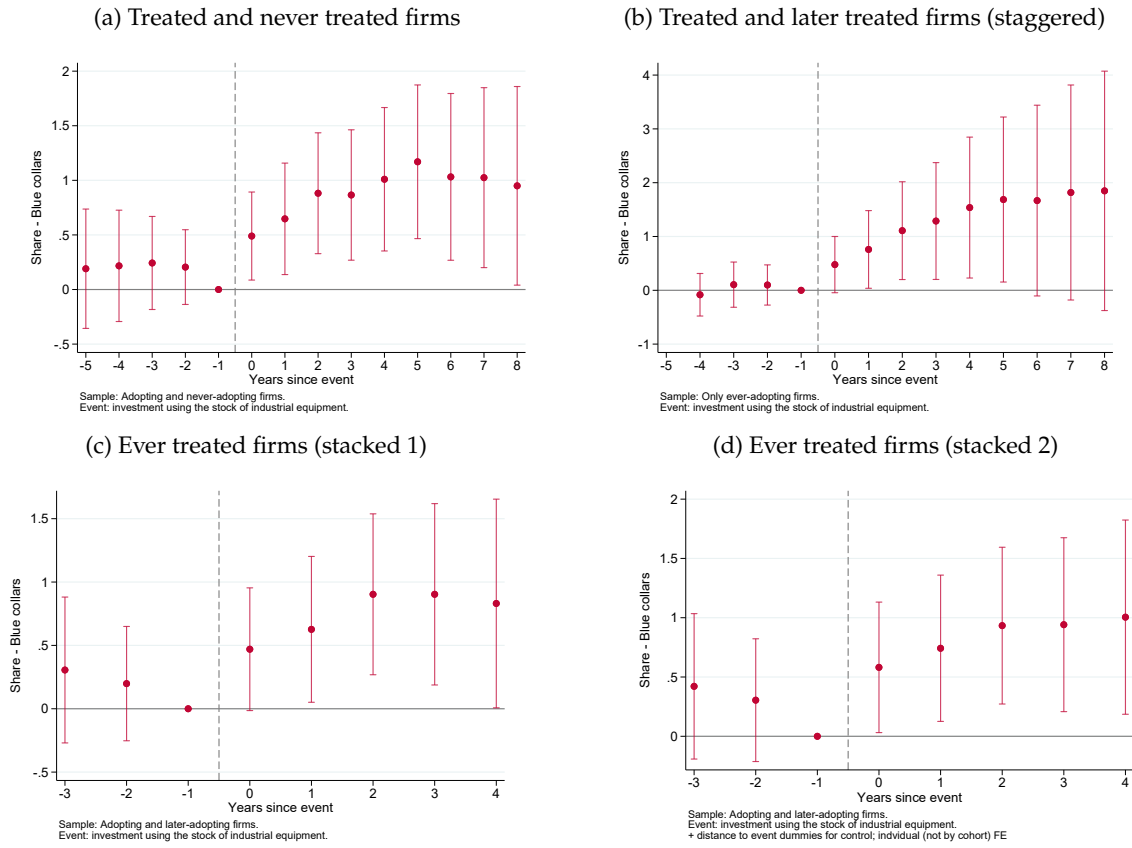
Notes: These figures plot the association of automation with the size of firms using different specifications. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Specification from: (a) Equation 2.1; (b) Equation 2.E.3; (c) Equation 2.E.4 but without the γ_k parameter; (d) Equation 2.E.4. The standard errors are clustered at the firm level.

Figure 2.E.8: Event studies at the firm level (results on sales) - comparison of different specifications



Notes: These figures plot the association of automation with sales using different specifications. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Specification from: (a) Equation 2.1; (b) Equation 2.E.3; (c) Equation 2.E.4 but without the γ_k parameter; (d) Equation 2.E.4. The standard errors are clustered at the firm level.

Figure 2.E.9: Event studies at the firm level (results on the share of blue-collar workers) - comparison of different specifications

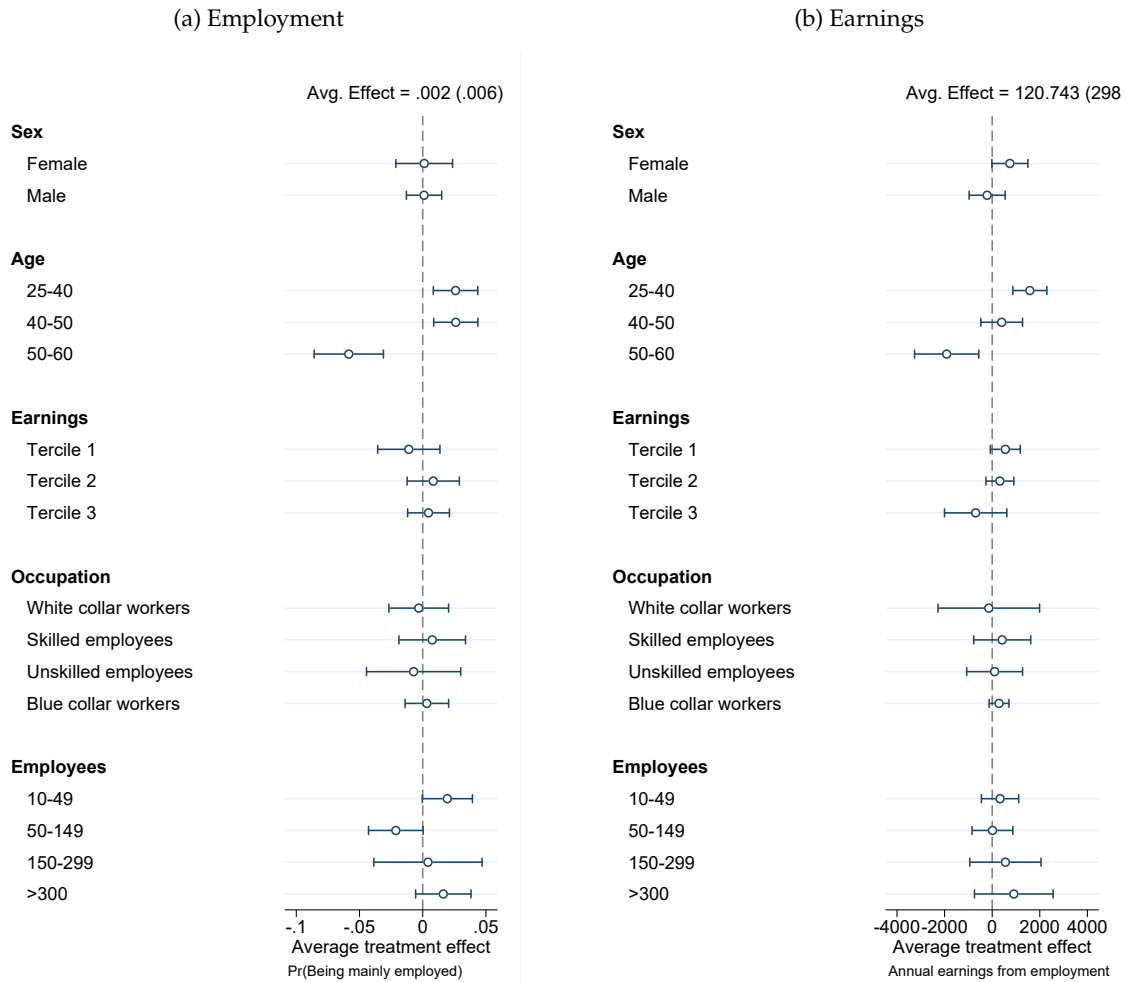


Notes: These figures plot the association of automation with the share of blue-collar workers using different specifications. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regressions. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The analyses are conducted on the narrowed sample of firms. The treated group is composed of all firms having experienced an automation event. The control group is made of never-treated firms. Year 0 denotes the year of the automation event. Specification from: (a) Equation 2.1; (b) Equation 2.E.3; (c) Equation 2.E.4 but without the γ_k parameter; (d) Equation 2.E.4. The standard errors are clustered at the firm level.

2.F Additional graphs for the Worker-level results

2.F.1 Labor market outcomes

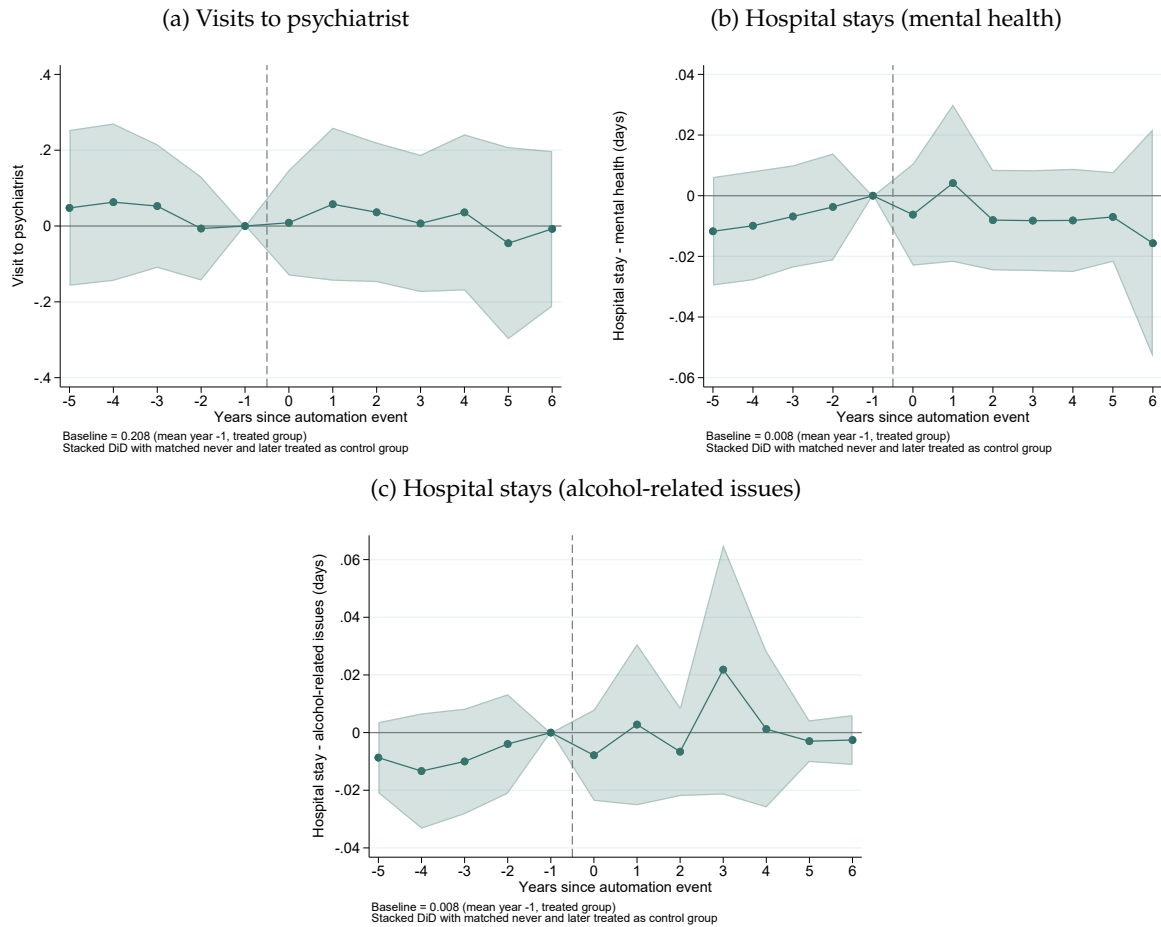
Figure 2.F.1: Heterogeneity (labor market outcomes)



Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 2.3). δ^j measures the difference in the evolution of the outcome (between the pre and post-period) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Outcomes: (a) P(Employment); (b) Annual earnings. The standard errors are clustered at the worker level.

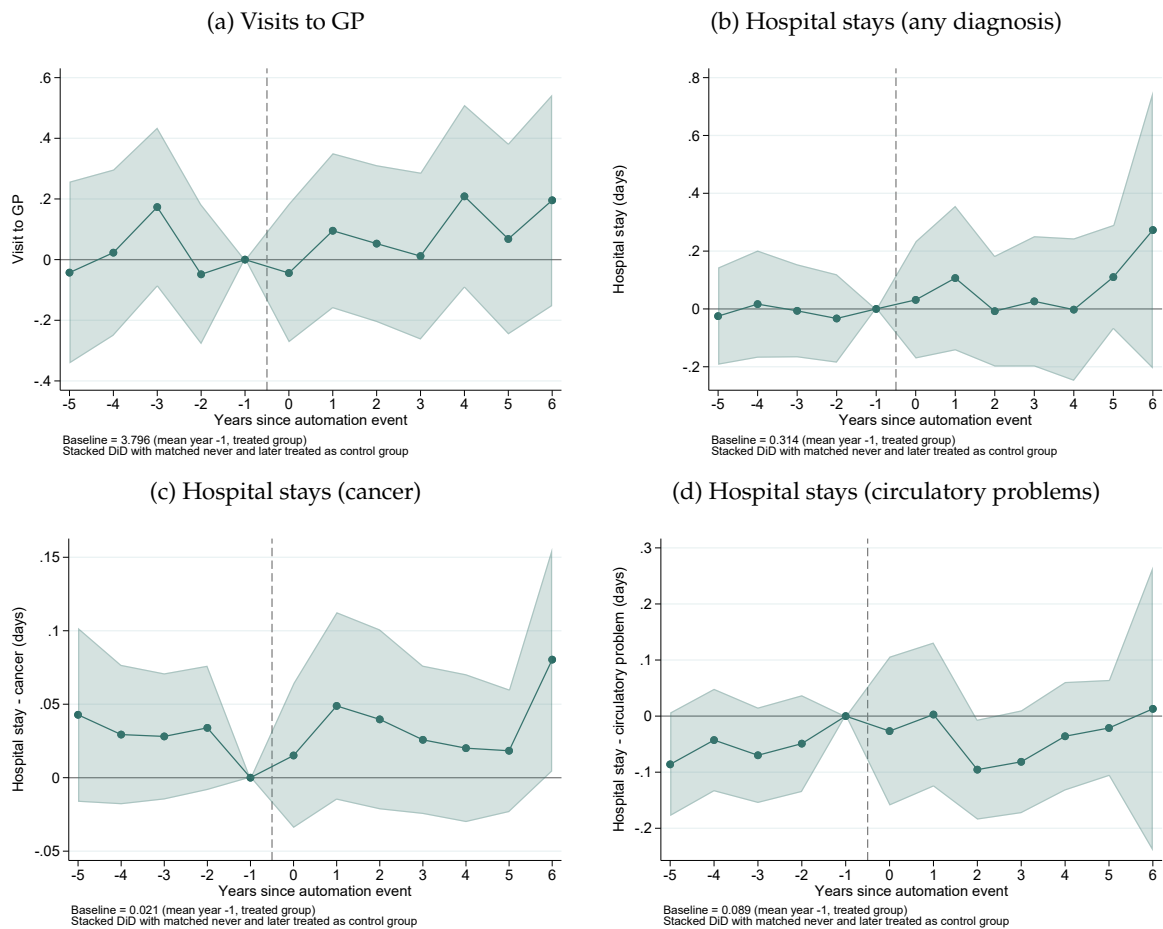
2.F.2 Health outcomes

Figure 2.F.2: Event studies at the worker level - mental health outcomes



Notes: These figures plot the association of automation with workers' mental health outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) Number of visits to a psychiatrist; (b) Length of hospital stays for mental health issues; (c) Length of hospital stays for alcohol-related issues.

Figure 2.F.3: Event studies at the worker level - physical health outcomes



Notes: These figures plot the association of automation with workers' mental health outcomes. These figures plot the δ_k 's coefficients obtained from the difference-in-difference regression of equation 2.2. δ_k measures the difference in the evolution of the outcome (between the relative year k and the reference year -1) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Year 0 denotes the year of the automation event. Outcomes: (a) Number of visits to a GP; (b) Length of hospital stays for any diagnosis; (c) Length of hospital stays for cancer; (d) Length of hospital stays for circulatory problems.

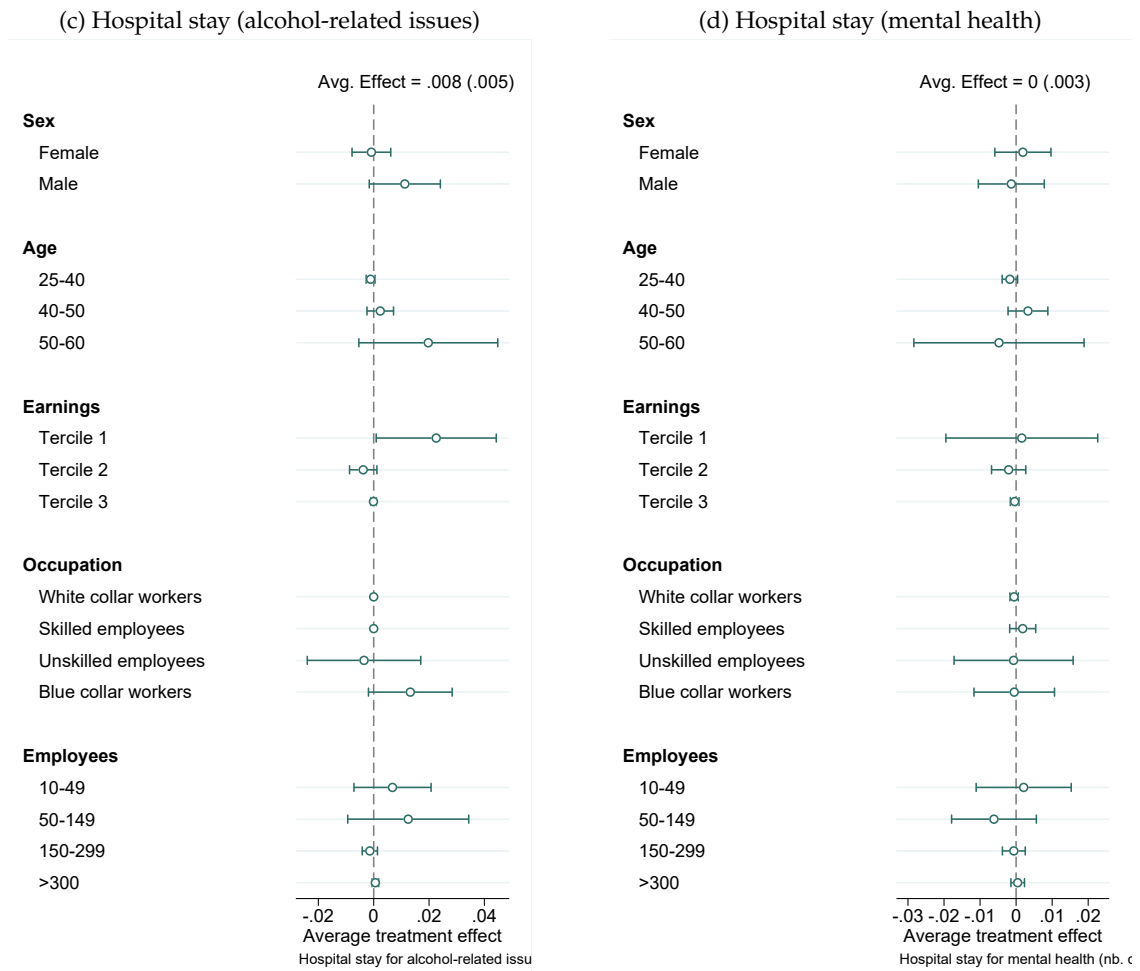
Figure 2.F.4: Heterogeneity (mental health outcomes)

(a) Sleeping pills



(b) Visit to psychiatrist





Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 2.3). δ^j measures the difference in the evolution of the outcome (between the pre and post-period) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Outcomes: (a) Number of boxes of sleeping pills; (b) Number of visits to psychiatrists; (c) Length of hospital stays for alcohol-related issues; (d) Length of hospital stays for mental health issues. The standard errors are clustered at the worker level.

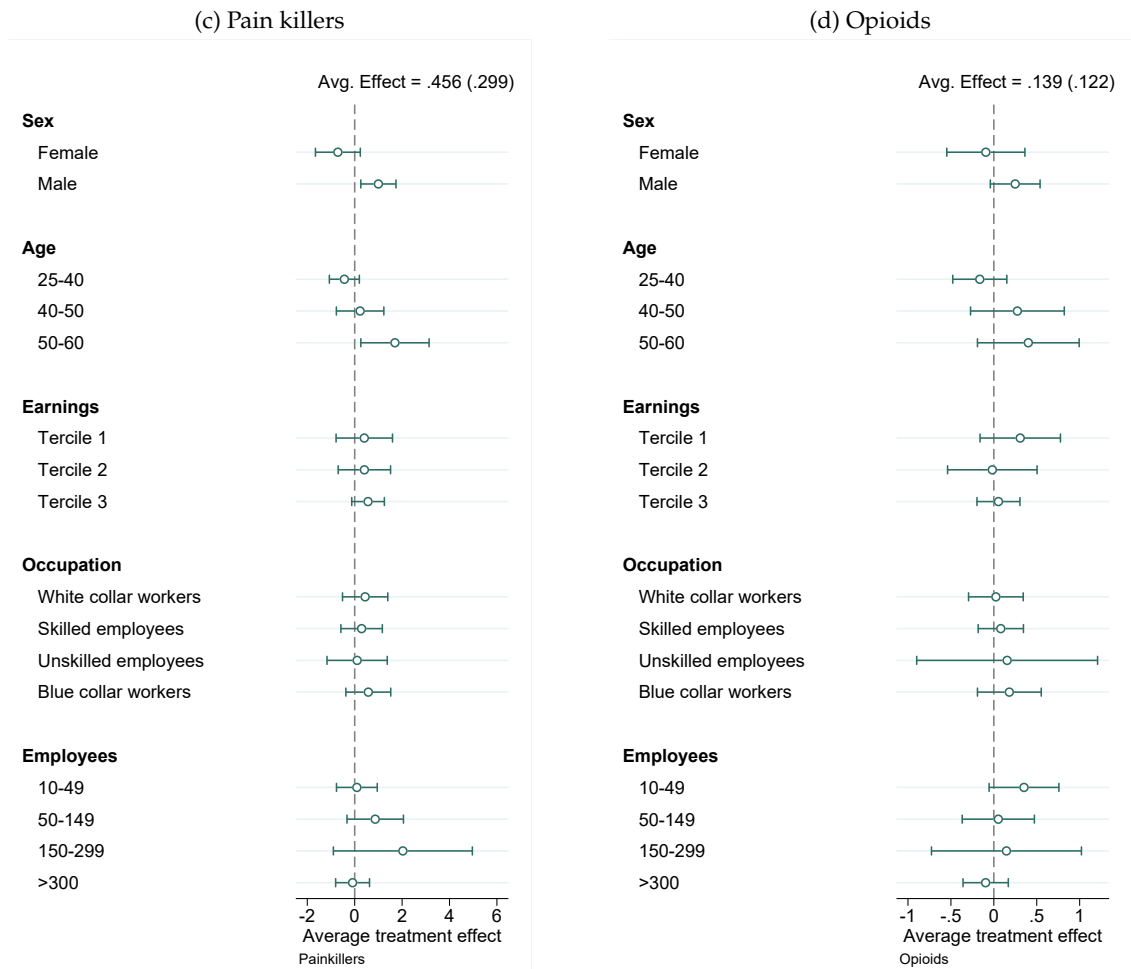
Figure 2.F.5: Heterogeneity (general health outcomes)

(a) Visits to GP



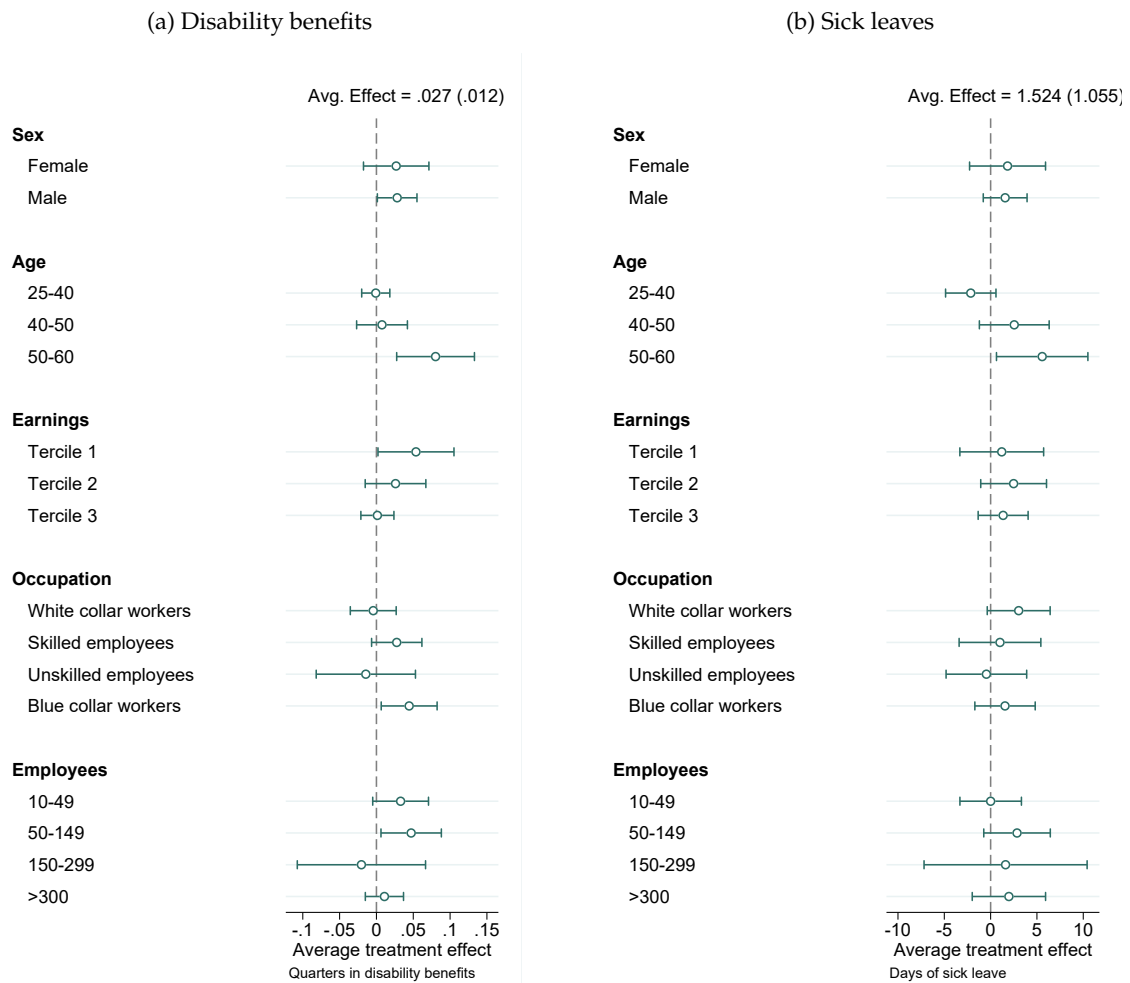
(b) Hospital stay (any diagnosis)





Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 2.3). δ^j measures the difference in the evolution of the outcome (between the pre and post-period) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Outcomes: (a) Number of visits to a GP; (b) Length of hospital stay for any diagnosis; (c) Number of boxes of pain killers; (d) Number of boxes of opioids. The standard errors are clustered at the worker level.

Figure 2.F.6: Heterogeneity (health insurance)



Notes: These figures plot the post-event coefficient (δ^j) following separate static event-study regressions (equation 2.3). δ^j measures the difference in the evolution of the outcome (between the pre and post-period) between the treated group and the control group. The treated group consists of workers exposed to a firm's automation event (investment in industrial equipment and machines). The control group is composed of matched workers who were treated later or never, with similar pre-event characteristics. Outcomes: (a) Number of quarters in disability benefits; (b) Number of days in sick leave. The standard errors are clustered at the worker level.

Chapter 3

Are Sugar Taxes Well Designed?

Empirical evidence from the UK soft drink market

This chapter is co-authored with Olivier Allais, Céline Bonnet and Maxime Tranchard.

Abstract

Policymakers have increasingly implemented nutritional taxes to influence consumer behavior toward healthier diets, with the efficiency of these taxes largely depending on their design. While theoretical and empirical literature suggests that taxes should be proportional to the harm caused, most nutritional taxes implemented to date, including sugar taxes, feature tiered rather than linear designs. This inconsistency between theory and practice raises the question: can tiered taxes be optimal, and if so, how should their key components be set? In this paper, we evaluate the performance of tiered sugar tax designs compared to the theoretically optimal linear tax. Using a welfare maximization framework, we account for externalities from excess sugar consumption, heterogeneous consumption patterns, and firms' strategic pricing behavior. Our findings reveal that tiered tax designs, when incorporating strategic responses from firms, can lead to significantly greater welfare improvements than the implemented Soft Drinks Industry Levy in the UK. Specifically, the optimal design features higher taxes on high-sugar products, prompting firms to reduce prices on lower-sugar alternatives. This adjustment not only enhances public health outcomes but also increases consumer surplus and preserves firm profitability. These results suggest that policymakers should carefully account for firms' strategic behavior when designing sugar taxation policies to maximize welfare benefits.

Keywords: Tax design, Welfare, Sugar tax

JEL classification: D62, H21, H30, I18, Q18

1 Introduction

In recent decades, policymakers have implemented "sin taxes" primarily to steer consumer choices toward healthier food options by targeting products considered harmful to society and individuals, such as alcohol, tobacco, and soft drinks. The taxes aim to reduce demand by increasing their prices and generate revenue for the government. To date, more than 50 countries have implemented taxes on sugar-sweetened beverages (SSBs).¹ However, the design of these taxes varies significantly from one country to another. The majority of designs implemented by regulators in the last 20 years include some tiered aspect (*e.g.*, France, UK, Spain, Chile). A notable aspect is that there has not been a sugar tax implemented that is strictly proportional to sugar content,² although it was theoretically and empirically shown it is the optimal design [Allcott et al., 2019a; O'Connell and Smith, 2024].

Can these tiered sugar-based taxes be optimal and if so, how should the main components of the taxes be set? Should all products be taxed, or only those with high sugar content? If it is the latter, which and how many sugar tax thresholds should be set? Is the level of taxation properly calibrated in relation to the social costs of sugar consumption? This paper aims to address the questions by analyzing how crucial components of tax design affect total welfare. Specifically, we propose a framework where the regulator maximizes the total welfare, integrating the social cost induced by excess sugar consumption. We account for the negative externalities resulting in healthcare costs associated with treating conditions caused by sugar consumption, such as weight gain, type 2 diabetes, and cardiovascular disease [Allcott et al., 2019b; Bhattacharya and Sood, 2011]. Including the cost of externalities allows for a comprehensive measure of welfare that accounts for heterogeneous effects on consumers, firms, the government, and health. Three components of the tax design can vary: the number of thresholds, the location of these thresholds (in grams of sugar per 100 mL), and the tax rate for each threshold (in euro cents per liter). This setting enables us to cover a large variety of tax

¹Obesity Evidence Hub [2022] shows a list of implemented policies and different tax designs.

²Quasi linear sugar taxes have been implemented but all have a threshold. For example, taxes at a rate of approximately 0.15 and 0.40 US cents for each gram of sugar over an initial threshold of 4 g/100 mL have been implemented in South Africa and Sri Lanka, respectively.

designs where the tax rate increases with the sugar content, from taxes proportional to sugar content to tiered tax designs with varying numbers of thresholds.

To achieve this objective, we focus on the UK market of non-alcoholic beverages, where the consumption of soft drinks has emerged as a notable health concern, prompting the government to enact the Soft Drinks Industry Levy (SDIL) tax in 2018.³ Sugar-sweetened beverages are major contributors to high sugar intake levels within the population (22% for children, 33% for adolescents and 21% for adults, [Public Health England \[2018\]](#)). We use rich scanner data obtained from a representative panel of UK households in 2017. Daily household purchases of differentiated products are provided in this dataset.

Our methodology establishes a regulatory framework that accounts for total welfare, measured as consumer surplus, firm profits, tax revenues, and the costs of externalities resulting from excessive sugar consumption. In order to estimate the four components of welfare, the methodology employed is based on the classic counterfactual evaluation in structural industrial organization from the seminal papers of [Berry et al. \[1995, 2004\]](#). We rely on the three-step structural econometric strategy approach used by [Bonnet and Réquillart \[2013\]](#). In the first step, we estimate a discrete choice model of demand allowing for substitutions both between and within varieties of non-alcoholic beverages. This estimation approach identifies household-specific preference parameters and the demand curves for the non-alcoholic beverage market. In the second step, we model the supply side as an oligopoly proposing differentiated products and competing *à la Nash* in a Bertrand game, in the spirit of [Berry et al. \[1995\]](#) and [Nevo \[2001\]](#). We use the estimated demand curves to identify the price-cost margins for each product and the unit costs of production for firms. Finally, the third step is the simulation of the counterfactual.

This paper contributes to the existing literature that focuses on the optimal design of taxes. The recent literature [[Allcott et al., 2019a](#); [O'Connell and Smith, 2024](#)] extend the theoretical results of [Pigou \[1934\]](#) and [Diamond \[1975\]](#) on the optimality of taxation

³The Soft Drinks Industry Levy is designed such that beverages with more than 8 grams of sugar per 100 mL are taxed at a high rate (24 pounds per liter), while those containing 5 to 8 grams of sugar per 100 mL face a lower tax rate (18 pounds per liter). Drinks with less than 5 grams of sugar per 100 mL are exempt. Additionally, pure fruit juices and beverages with high milk content are not subject to the tax.

at the marginal cost of the externality to the particular case of "sin taxes". [Allcott et al. \[2019a\]](#) show that the optimal tax on sugar-sweetened beverages should be based on sugar content and that volumetric taxation is less efficient. [O'Connell and Smith \[2024\]](#) validate these results in case of market power. In our study, we extend their analyses to all existing designs where sugar is taxed, including tiered sugar-based tax.

Our paper is also related to [Griffith et al. \[2019\]](#), who discuss the optimal multiple tax system for alcoholic beverages. They show that in the presence of nonlinear externalities, when heavy consumers differ in their consumption patterns, multi-rate taxes are optimally more efficient than single-rate taxes. Our approach does not focus on the functional form implications of the externality, but aims to determine the optimal number of thresholds, their location, and the optimal tax levels, considering the heterogeneity in consumption patterns. We show that tiered taxes are more efficient even without nonlinear externality in the case of strategic firms.

Finally, our paper is related to the large literature on nutrition policy evaluation, especially on soft drink sugar taxes with empirical structural estimates. In particular, [Dubois et al. \[2020\]](#) show that the soft drink tiered sugar-based tax implemented in the UK partially failed to target the groups whose behavior the regulator would most like to influence. Our framework, which integrates the social cost of sugar and accounts for heterogeneity in demand, shows that tiered sugar-based taxes can be optimal.

Our findings outline the optimal tax structure for two scenarios. In a single-threshold design, items with sugar content above 9 grams per 100 mL are subject to a flat tax of 3.05 € per liter, while those below this threshold are exempt. In a two-threshold design, products containing between 2 and 9 grams of sugar per 100 mL incur a tax of 0.15 € per liter, and those exceeding 9 grams per 100 mL face a higher tax of 5.23 € per liter. Our results show that optimal sugar tax designs which account for firms' strategic pricing lead to significantly greater welfare improvements than designs ignoring firms' strategic reactions such as linear tax, or the implemented SDIL. The optimal tax design, through higher taxes on high-sugar products, creates a new equilibrium where firms lower prices on low-sugar alternatives, resulting in improved public health outcomes and consumer surplus. Additionally, firms are able to preserve profits through strate-

gic price adjustments. Policymakers seeking to maximize the welfare impact of sugar taxes should, therefore, carefully consider firms' strategic responses when designing their taxation policies.

The paper is organized as follows. Section 2 describes the UK non-alcoholic beverage market. Section 3 details our methodology. Section 4 provides the main results. Section 5 concludes.

2 The UK soft drink market

2.1 Data

We use representative consumer panel data from Kantar Worldpanel in the United Kingdom for 2017. These data are collected by the market research firm Kantar Worldpanel and are widely used in the economic literature on industrial organization and food purchases in Europe [Dubois et al., 2022]. The dataset covers 24,586 households who are representative of the British population, and includes detailed information on household members, such as age, socio-economic status, income class, and weight status.⁴ Households record their daily grocery purchases brought into the home by scanning the bar codes of each item. Thus, the data provide high-frequency details on disaggregated products, including product attributes (*e.g.*, brand, size, sugar content), store of purchase, quantity, and price. This analysis focuses on purchases of non-alcoholic beverages for the year 2017.⁵

Demographic characteristics. We categorize households according to their composition (number and age of children), obesity status (proportion of obese or overweight adults), and socio-economic class (see Appendix 3.A.1 for details and Table 3.A.1 for summary statistics).

Market & product definition. The non-alcoholic beverage market includes sugar-sweetened beverages (SSBs), fruit juices, flavored waters, and milk-based drinks.⁶ We define products by crossing information on the firm (*e.g.*, Coca-Cola Company), the

⁴Additionally, Kantar Worldpanel provides weights to ensure the panel's representativeness and to correct for reporting biases, including periods when households are away from home.

⁵Note that the SDIL had not yet been implemented in the UK. We assume that consumer preferences remain stable over time.

⁶We exclude water and products that are not ready-to-drink (*e.g.*, syrup, powdered drinks, cocktail mixers). See Table 3.A.2 for a description of the products included in the market.

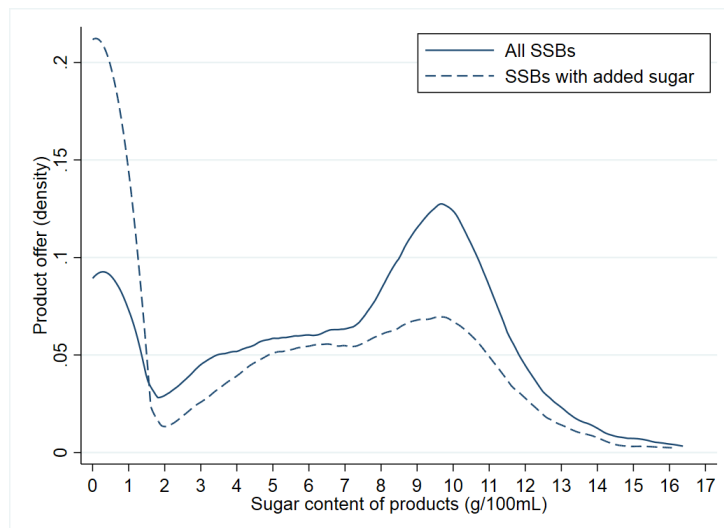
brand (e.g., Coca-Cola, Fanta, Sprite, ...), the 11 drinks sub-categories,⁷ and other product characteristics (diet or regular characteristic for SSBs; presence of added sugar and flavor for fruit juices). We get a set of 402 distinct varieties of beverages (hereafter, called alternatives) that we consider in our analysis. Additional details are provided in Appendix 3.A.3.

Prices. The price of each alternative is calculated as the weighted ratio of total expenditure in euros over the total quantities in liters of all items belonging to the alternative in the corresponding four-week period. The weights used are Kantar period-specific household sample weights.

2.2 Descriptive statistics

Sugar distribution. The distribution of sugar in the products offered on this market shows two peaks: one for products containing no sugar and one for products containing around 10 grams of sugar per 100ml (Figure 3.1). 34% of products have less than 5g of sugar per 100 mL, 20% of products have between 5g and 8g of sugar per 100 mL and 46% of products have more than 8g of sugar per 100 mL.

Figure 3.1: Distribution of the sugar content of the non-alcoholic beverage supply



Notes: This figure plots the distribution of sugar and added sugar content (in grams per 100mL) of the supply of products.

Price. Non-alcoholic beverages have an average price of 0.67 € (Table 3.A.3). Products with added sugar are on average less expensive. The most expensive sub-categories

⁷See Table 3.A.4 for descriptive statistics about the different subcategories.

are smoothies and energy drinks. There is no unequivocal pattern between price and diet characteristics: depending on the sub-category, diet products can be more or less expensive than regular products (Table 3.A.4).

Purchases. Households purchase on average 65 liters per capita per year, among which 44 liters contain added sugar (Table 3.A.5). Only 13% of non-alcoholic beverages purchased by households have a sugar content above 8 g per 100 mL. A large proportion of non-alcoholic beverages purchased by households are beverages with a sugar content between 0 and 5 g per 100 mL. 14.5% of Kantar households (7.9% when considering only products containing added sugar) purchase more than the equivalent of one can (330 mL) per day per capita (Figures 3.A.2a and 3.A.2b). 33.6% of households purchase more than one sugar cube per day per capita (Figure 3.A.3).

3 Methodology

3.1 Welfare maximization program

3.1.1 Definition of a tax design

We consider that a tax design is defined by three parameters that the regulator has to pick: i) the number k of thresholds, ii) a vector ρ of dimension k with the location of thresholds (in g/100mL), and iii) a vector ω of dimension k with the tax level for each threshold (in € cents/L). The vector $\tau(k, \rho, \omega)$ of dimension J contains the resulting tax amount supported by each product j for the tax design defined by (k, ρ, ω) .⁸

This definition encompasses all tax designs enacted in various countries and states where the tax rate increases with the sugar content, such as a volumetric tax (by setting the number of thresholds to 1 and the location of the threshold at 0 g/100mL, all products are taxed at a single tax rate, which is equivalent to a volumetric tax) and a linear tax or a tiered tax with proportional tax rate within each tier (by proposing a number of thresholds equal to the number of distinct sugar values in the market).⁹

⁸We will implement the tax on the same set of products as in the Soft Drink Industry Levy, *i.e.*, all soft drinks except for pure fruit juices and drinks with high milk content.

⁹We simplify the optimization program of equation 3.11 by restricting the sugar content to integer values. This approach reduces the dimensionality of the problem, resulting in 17 discrete sugar levels for the products in our market. Figure 3.A.1 displays the distribution of the sugar content of products (supply dimension) when discretizing the sugar content.

3.1.2 Theoretical welfare definition

We provide a conceptual framework where a regulator wants to implement a sugar tax. The regulator's objective is to find a design that maximizes total welfare. The welfare encompasses consumers' utility, firms' profits, tax revenue, and the external costs induced by excess sugar consumption, as follows:

$$\text{Welfare} = \text{Consumer surplus} + \text{Firm profit} + \text{Fiscal revenue} - \text{Social cost of sugar} \quad (3.1)$$

Consumer surplus. The consumer surplus reflects the benefits consumers derive from purchasing a product. It is measured as the difference between what they are willing to pay and the price they actually pay.¹⁰ Consumer surplus is determined by consumers' preferences for product attributes (*e.g.*, price sensitivity, taste for sugar, or brand loyalty). To quantify these preferences, we will use a demand model (section 3.2.1). Given that the introduction of the tax will likely affect product prices, we need to understand how consumers will react and this depends on their valuation of product attributes. This model will also allow us to estimate price elasticities, which capture how sensitive consumers are to price changes. Both own-price and cross-price elasticities will be required to account for substitution patterns among available alternatives.

Firm profit. A firm's profit is the amount it earns after selling a product, computed as the difference between total revenue and total costs. It depends on the product's price, marginal costs, and the quantity sold. The implementation of a tax will mechanically increase the marginal cost of each product targeted by the tax. In response, firms may decide to adjust their product prices.¹¹ They might choose to absorb part of the tax, resulting in a price increase that is less than the tax amount, or increase their margins, leading to a price rise that exceeds the tax amount. We will introduce a supply model (section 3.2.2) to capture firms' strategic responses to the tax and calculate the tax pass-through, which indicates how much of the tax is passed on to consumer prices. The new profit will depend on a combination of the adjusted prices and the resulting quan-

¹⁰McFadden [1999] formalized the change in consumer surplus with the "random compensating variation" approach, *i.e.*, the income adjustment required to equalize maximum utility with the tax and without. In other words, how much compensation (in monetary terms) would consumers require for the introduction of the tax to maintain the same level of maximum utility?

¹¹Firms may also respond by modifying their offer of products (introducing new products or reformulating existing ones). However, we exclude these adjustments from our analysis.

tities sold (which in turn will be determined by consumer demand in response to these price changes).

Fiscal revenue. Fiscal revenue refers to the tax income collected by the government. It is primarily driven by demand factors, as it depends on the quantities purchased by consumers. Generating revenue is not the primary intent of the tax, but the funds raised are often allocated to public health programs or other societal benefits.

Social cost of sugar. The social cost of sugar refers to the negative externalities resulting from excess sugar consumption. High levels of sugar intake contribute to various health issues, leading to increased healthcare expenditures for treating sugar-related conditions. This qualifies as an externality because the costs of healthcare are borne by society as a whole, rather than solely by the individuals who consume excessive sugar [Allcott et al., 2019b; Bhattacharya and Sood, 2011]. The social cost of sugar is affected by demand factors, as it depends on consumers' choices of products and the quantities they decide to purchase. We obtain the social cost of sugar from the estimate provided by Rischbieth et al. [2020], which is 27€ per kg of sugar. For sensitivity analyses, we consider a lower bound at 10€ per kg of sugar.¹²

3.2 Consumer and firms behaviors

We represent consumers' and firms' behavior using a flexible discrete-choice model of demand for differentiated products with a supply model assuming an oligopolistic competition. The estimation method is in two steps. First, we estimate a demand model in order to analyze household preferences for non-alcoholic beverage purchases. Second, using the estimated consumers' substitution patterns, we model the supply side to determine pricing strategies and alternatives' marginal costs.

3.2.1 Demand model

We consider a flexible discrete-choice model to estimate the demand and obtain the price elasticities for every alternative. Specifically, we use a random coefficient logit model (RCLM) [Berry et al., 1995; McFadden and Train, 2000]. In this model, prefer-

¹²See section 4.2.2 for a discussion on the implications of the choice of the value of the healthcare externality.

ences over product characteristics are specified in a flexible manner, as it allows for both observed and unobserved heterogeneity in the slopes of the utility function and allow for flexible consumer substitution patterns.

Following [Revelt and Train \[1998\]](#), let t denote the index of time ($t = 1, \dots, T$), i the index of the household in the sample ($i = 1, \dots, N$) and j the index of the product inside the choice set of differentiated products ($j = 1, \dots, J_t$) at time t .¹³

Utility. The indirect utility function V_{ijt} for household i buying product j in period t is given by

$$V_{ijt} = \alpha_i p_{jt} + \beta_i X_{jt} + \epsilon_{ijt} \quad (3.2)$$

where p_{jt} is the price of product j in period t , α_i is the marginal disutility of the price for household i , X_{jt} is a vector of observed product characteristics, β_i is the vector of associated parameters that capture the individual taste for product characteristics, and ϵ_{ijt} is an individual error term.

We assume that parameter α_i varies across households, with some unobserved heterogeneity. Indeed, households can have a different price disutility. It can be rewritten as

$$\alpha_i = \alpha + \pi^\alpha D_i + \sigma v_i \quad (3.3)$$

where α is the mean marginal disutility of the price for all households, π the vector of parameters associated with demographic characteristics D_i (family composition, obesity status and socio-economic class of the household), and v_i measures the unobserved heterogeneity of the households. We denote $P_v(\cdot)$ the distribution of parameter v .

Additionally, the parameter β_i also depends on household characteristics. It can be rewritten as $\beta_i = \beta + \pi^\beta H_i$ where β includes brand and category fixed effects and π^β are the coefficients associated with a dummy for diet products and the sugar content of fruit juices and SSBs, interacted with household composition characteristics H_i .

We can divide the indirect utility between a mean utility $\delta_{jt} = \alpha p_{jt} + \beta X_{jt} + \xi_{jt}$ where ξ_{jt} captures all unobserved product characteristics and a deviation from this mean utility

¹³Here t corresponds to a four-week period ($t = 1, \dots, 13$).

$\mu_{ijt} = (\sigma v_i + \pi^\alpha D_i)p_{jt} + \pi^\beta H_i X_{jt}$. Hence the indirect utility is given by

$$V_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt} \quad (3.4)$$

A household is defined by the vector $(v_i, D_i, \epsilon_{i0t}, \dots, \epsilon_{ijt})$.

Outside option. The household can decide not to buy any of the considered products. The utility of this option is normalized to zero. The indirect utility of choosing the outside option is written as $V_{i0t} = \epsilon_{i0t}$.

Market share. We assume that ϵ_{ijt} is independently and identically distributed as an extreme value type I distribution. The conditional probability that household i chooses product j in period t is:

$$s_{ijt}(v) = \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^J \exp(\delta_{kt} + \mu_{ikt})} \quad (3.5)$$

The market share of product j in period t is [Nevo, 2001]:

$$s_{jt} = \int_{A_{jt}} \left(\frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^J \exp(\delta_{kt} + \mu_{ikt})} \right) dP_v(v) \quad (3.6)$$

where A_{jt} is the set of households who have the highest utility for product j in period t , and P_v is the cumulative distribution function of v which is typically assumed to be standard normal.

Elasticity. The random coefficient logit model generates a flexible pattern of substitutions between products. We can then write the own-price and cross-price elasticities of the market share s_{jt} as:

$$\theta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} \frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt} (1 - s_{ijt}) \phi(v_i) dv_i & \text{if } j = k \\ -\frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} \phi(v_i) dv_i & \text{otherwise} \end{cases} \quad (3.7)$$

where $\phi(\cdot)$ is the density function of $P_v(\cdot)$.

Identification. This method relies on the assumption that all product characteristics X_{jt} are independent of the error term ϵ_{ijt} , which can be decomposed into a product-specific error term and an individual error term, $\epsilon_{ijt} = \xi_{jt} + e_{ijt}$. However, there is

empirical evidence that unobserved factors included in ξ_{jt} can be correlated with observed characteristics X_{jt} , producing endogeneity problems [Berry, 1994]. Such unobserved characteristics can be promotions or advertising, for example. For instance, promoted products are often moved to the front of the shelf, advertised, and sold at a lower price at the same time. Since we do not have any information on advertising expenditure spent by firms, the estimated impact of observed prices on demand will then capture both a true price effect and the effect of unobserved marketing efforts. Prices may also be endogenous if some unobserved characteristics are positively valued by consumers, who are thus ready to pay a premium for them.

We use a control function approach as in Petrin and Train [2010] to account for price endogeneity. We use the price index for glass which is used in the production of non-alcoholic beverages.¹⁴ This instrument is not likely to be correlated with unobserved determinants of demand for non-alcoholic beverages because this industry only represents a very small share of the demand for those inputs. We also have two BLP instruments: the number of competing products offered by other firms within the product category and the total sugar content of competing products within the sugar-based category.¹⁵ The instruments are not weak, the F-test for the joint test of the nullity of the instrumental variables is superior to 10 (Table 3.B.1).

Estimation. We estimate the demand model using the simulated maximum likelihood method as in Revelt and Train [1998]. Details are provided in section 3.B.2.

3.2.2 Supply model

We consider F firms that compete in prices on the considered market, sell products to consumers and set prices. At each period, the firm maximizes its profit, conditional on the demand parameters and other firms' prices, holding the set of products offered

¹⁴Cost shifters are Office for National Statistics price production indices (<https://www.ons.gov.uk/economy/inflationandpriceindices>).

¹⁵The nutritional category refers to the regular or diet characteristic for SSBs and milk-based drinks and to the three levels of sugar content for fruit juices "Low sugar-sweet", "Sugar-sweet" and "High sugar-sweet" (these three sugar levels are based on the terciles of the sugar content distribution and are computed for each category separately).

and every other observed and unobserved characteristics constant:

$$\Pi_{ft} = \sum_{j \in G_{ft}} [M_t(p_{jt} - c_{jt})s_{jt}(p_t)] \quad (3.8)$$

where G_{ft} is the set of products sold by firm f in period t , M_t is the size of the market in period t , p_{jt} is the price of product j in period t , c_{jt} is the constant marginal cost to produce and sell product j in period t , $s_{jt}(p_t)$ is the market share of product j in period t given the vector of product price p_t .

The equation of prices of products $j \in G_{ft}$ is determined with the first order conditions:

$$s_{jt}(p_t) + \sum_{k \in G_{ft}} [M_t(p_{kt} - c_{kt}) \frac{\partial s_{kt}}{\partial p_{jt}}] = 0 \quad \forall j \in G_{ft} \quad (3.9)$$

We recover estimates of firms' margins $\gamma_{jt} = p_{jt} - c_{jt}$ for each product using the first-order conditions and estimates of the demand model. Using equation (3.9), the vector of margins $\gamma_t = (p - c)_t$ can be written in matrix notation.

$$\gamma_t = (p - c)_t = - \left(\sum_{f=1}^F I_{ft} S_{pt} I_{ft} \right)^{-1} \left(\sum_{f=1}^F I_{ft} s_t(p_t) \right) \quad (3.10)$$

where I_{ft} is the ownership diagonal matrix ($J_t \times J_t$) of firm f in period t with elements $I_{ft}(j, j)$ equal to 1 if product j is produced by firm f in period t and zero otherwise, S_{pt} is the matrix ($J_t \times J_t$) of the first derivatives of all market shares with respect to all prices in period t , *i.e.*, $S_{pt} = \left(\frac{\partial s_{kt}}{\partial p_{jt}} \right)_{(j=1, \dots, J_t; k=1, \dots, J_t)}$ and $s_t(p_t)$ is the vector of product market shares in period t . We then derive estimates of marginal costs, given observed prices.

3.2.3 Results of the model

We present the main results of the demand model, price elasticities and firms' margins in this section.

Demand estimates. The estimates from the random coefficient logit model are presented in Table 3.B.2. Price has a significant and negative impact on utility for all populations. Price unobserved heterogeneity is also substantial. Households from the

poor and average classes are more sensitive to price than rich households. Households prefer diet products to regular products. Households have high brand loyalty and the choice of the brand prevails over the taste for sugar (brand fixed effect estimates, not displayed in the Table, ranging from 0.05 to 14.2 in absolute terms, are large compared to the other preferences parameter estimates).

Elasticities and demographic characteristics. When comparing the elasticities by household characteristics, we found that the demand for non-alcoholic beverages is the most elastic for households with 7-16 years old children, households with all adults overweight or obese and poor households (Table 3.B.3). We find that the sweeter the non-alcoholic beverage, the more elastic the demand for products; and it is also interesting to note that the elasticities of non-alcoholic beverages with a sugar content above 10 g per 100 mL remain almost constant. (Table 3.B.4). The results are valid for all household characteristics considered. Furthermore, all previous results in Table 3.B.3 are still valid for all sugar content considered.

Supply. The estimated margins and own-price elasticities are presented by firm in Table 3.C.1 and by sugar content in Table 3.C.2. Margins, expressed as a percentage of price, vary from 8.7% to 28.4% for national brands and are higher for private labels (52.3%). Own-price elasticities, range from -3.5 to -12.5 for national brands and are quite low for private labels (-4.4). Firms with less price-sensitive demand (lower absolute elasticity) are able to maintain higher profit margins. The margins are quite similar across all sugar categories, ranging from 17.7% to 22.5%, except for the category of sugar-sweetened products with less than 5 grams of sugar, which has a significantly higher margin of 49.9%.

3.3 Counterfactual simulations

We now describe the methodology to obtain the optimal parameters of a tax design to enhance total welfare.

3.3.1 Objective function

Total welfare. The regulator wants to maximize the theoretical welfare defined in equation 3.1. In practice, maximizing the theoretical welfare is equivalent to maximizing the change in welfare induced by the implementation of a tax.^{16,17}

For a given number of thresholds k , we look for the optimal combination of thresholds location ρ^* and tax rates ω^* that maximize the increase in total welfare. In other words, for any fixed $k \in \mathbb{N}$, the optimal tax design $\tau^*(k, \rho^*, \omega^*)$ is obtained from the following maximization program.

$$(\rho^*, \omega^*) = \arg \max_{(\rho, \omega)} \Delta W_t(k, \rho, \omega) \quad (3.11)$$

where

$$\Delta W_t(k, \rho, \omega) = \Delta W_t^C(k, \rho, \omega) + \Delta W_t^\pi(k, \rho, \omega) + \Delta W_t^{\text{tax}}(k, \rho, \omega) + \Delta W_t^{SC}(k, \rho, \omega) \quad (3.12)$$

Welfare changes arise from potential price adjustments, driven by shifts in the marginal cost of products following the introduction of a tax. These new prices following the implementation of the tax (k, ρ, ω) are denoted $p_t^{\text{post}(k, \rho, \omega)}$. The method used to calculate the new price equilibrium is detailed in section 3.3.2.

Welfare components. The change in total welfare (equation 3.12) is the sum of the changes in consumer surplus (equation 3.13), firm profit (equation 3.14), fiscal revenue (equation 3.15) and externalities after the implementation of a tax (equation 3.16).¹⁸

Consumer surplus. Following [McFadden \[1981\]](#) and [Small and Rosen \[1981\]](#), we compute the change in consumer surplus as:

$$\Delta W_t^C(k, \rho, \omega) = M_t \int \frac{\ln[\sum_{j=0}^J \exp(V_{ijt}(p_t^{\text{post}(k, \rho, \omega)})^j)] - \ln[\sum_{j=0}^J \exp(V_{ijt}(p_t)^j)]}{\alpha_i} dP_v(v) dP_D(D) \quad (3.13)$$

where $V_{ijt}(p_t^{\text{post}(k, \rho, \omega)})$ is the utility where the consumer price $p_t^{\text{post}(k, \rho, \omega)}$ is adjusted by

¹⁶Indeed, the observed welfare is a constant so it is equivalent to maximizing the difference between the welfare post-tax and the observed welfare.

¹⁷Note that for computational reasons, we only run the counterfactual simulations for a fixed period of 4 weeks. We picked the first period and we will perform robustness checks on other periods.

¹⁸We give equal weights to each welfare component.

the introduction of the tax (k, ρ, ω) .

Firm profit. The change in firms' profit is defined by:

$$\Delta W_t^\pi(k, \rho, \omega) = M_t \left(\sum_{f=1}^F \left(\sum_{j \in G_{ft}} \gamma_{jt}(p_t^{post(k, \rho, \omega)}) s_{jt}(p_t^{post(k, \rho, \omega)}) \right) - \sum_{f=1}^F \left(\sum_{j \in G_{ft}} \gamma_{jt}(p_t) s_{jt}(p_t) \right) \right) \quad (3.14)$$

where $s_{jt}(p_t^{post(k, \rho, \omega)})$ and $\gamma_{jt}(p_t^{post(k, \rho, \omega)})$ are the market shares and margins respectively of product j after the implementation of the tax (k, ρ, ω) , $s_{jt}(p_t)$ and $\gamma_{jt}(p_t)$ are the estimated pre-tax market shares and margins of product j .

Tax revenue. The amount raised by the implementation of the sugar tax is defined as follows:¹⁹

$$\Delta W_t^{\text{tax}}(k, \rho, \omega) = M_t \left(\sum_{j=0}^J \tau_j \cdot s_{jt}(p_t^{post(k, \rho, \omega)}) \right) \quad (3.15)$$

where τ_j is the tax level of product j .

Externality. The change in the externality is computed as:

$$\Delta W_t^{SC}(k, \rho, \omega) = M_t \left(\sum_{j=0}^J \kappa \cdot (\lambda_{jt} \cdot s_{jt}(p_t^{post(k, \rho, \omega)}) - \lambda_{jt} \cdot s_{jt}(p_t)) \right) \quad (3.16)$$

where κ is the social cost of sugar (in euros per kilogram) and λ_{jt} is the quantity of sugar per kilogram of product j .

3.3.2 New price equilibrium

The introduction of a tax will change the marginal cost of each product, which will lead to a potential change in their price. From the firms' profit maximization program defined in section 3.2.2, we estimate a vector of marginal costs that we denote $\hat{c}_t = p_t - \gamma_t = (\hat{c}_{1t}, \dots, \hat{c}_{jt}, \dots, \hat{c}_{Jt})$. For any tax design (k, ρ, ω) , we retrieve a new cost vector $c_t^{(k, \rho, \omega)}$ by adding the tax cost to the estimated marginal cost ($c_t^{(k, \rho, \omega)} = \hat{c}_t + \tau$). We then find the new equilibrium prices vector in period t , denoted $p_t^{post(k, \rho, \omega)} =$

¹⁹The tax is not subject to VAT in the UK so we only consider the amount raised through the implementation of the sin tax.

$(p_{1t}^{post(k,\rho,\omega)}, \dots, p_{jt}^{post(k,\rho,\omega)})$, using the following optimizing program:

$$\min_{\{p_{jt}^{post(k,\rho,\omega)}\}_{j=1,\dots,J_t}} \left\| \underbrace{p_t^{post(k,\rho,\omega)} - \gamma(p_t^{post(k,\rho,\omega)})}_{\hat{c}_t(p_t^{post(k,\rho,\omega)})} - c_t^{(k,\rho,\omega)} \right\| \quad (3.17)$$

which minimizes the difference between the marginal cost derived from the new price equilibrium and the marginal cost estimated from the model and shocked by the tax.

4 Results

In this section, we provide evidence on the optimal design of tiered taxes. We derive and compare optimal designs for one-tiered and two-tiered taxes against linear taxes and the SDIL, accounting for firms' strategic price reactions.²⁰ Furthermore, we investigate the impact of accounting for firms' strategic price reaction in designing sugar taxes. Lastly, we conduct robustness checks.

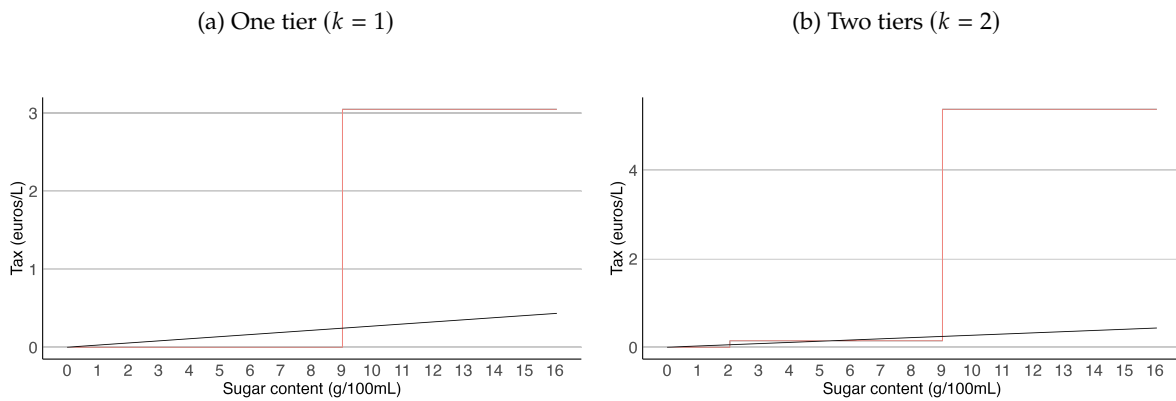
4.1 Optimal design of sugar taxes

Description of the designs. We obtain the optimal taxation scheme as a result of the regulator's maximization program (Figure 3.1). In the one-threshold design, products with sugar content exceeding 9 grams per 100 mL are taxed at a flat rate of 3.05 € per liter, while those below this threshold are exempt from the tax. In the two-threshold design, products with sugar content between 2 and 9 grams per 100 mL are taxed at 0.15 € per liter, whereas those with more than 9 grams per 100 mL face a higher tax of 5.23 € per liter.²¹

²⁰We limit the regulator's welfare maximization program to designs with one or two thresholds to ensure that our problem is concave, and for the sake of computation time. This limitation does not impede our methodology to enlighten tiered designs that perform largely better than the linear tax.

²¹This amount represents a significant share of the average price of SSBs. For now, we interpret it as a prohibitive cap on high sugar-sweetened products.

Figure 3.1: Comparison of tiered tax designs



Notes: The two figures plot the optimal tiered tax design with one and two thresholds (in red) where we account for firms' strategic behavior. The optimal tax rates and location of thresholds are obtained empirically from the maximization program presented in section 3.3.1. The black line corresponds to the theoretical optimal linear tax design where the tax rate is set at the value of the social cost (27 euros per kg of sugar). The y-axis plots the tax rate (in euros per L) and the x-axis displays the sugar content of products (in grams per 100 mL).

Welfare decomposition. To better understand the efficiency of these designs, we decompose the welfare effects into their key components: firms' profit, consumer surplus, tax revenue, and externalities. Table 3.1 presents the results of this decomposition, comparing the welfare effects of the one-threshold and two-threshold optimal designs with a linear tax design, which is identified as the optimal design in the theoretical literature.

The linear tax design leads to a total welfare gain of 5.64 million €, largely driven by a reduction in externalities (8.37 million €) and an increase in tax revenue (4.17 million €). However, the significant decrease in consumer surplus (-6.45 million €) and a slight reduction in firm profits (-0.45 million €) offset these gains.

In contrast, the one-threshold optimal tax design results in a much larger welfare gain of 10.11 million €. This is primarily driven by an increase in consumer surplus (+3.06 million €) and a substantial reduction in externalities (6.64 million €). Firm profits also see a modest positive increase (+0.41 million €), although there is no net change in tax revenue.

The two-threshold design yields the highest total welfare improvement, amounting to 21.44 million €. This design achieves the largest externality reduction (10.57 million €) while significantly increasing consumer surplus (+7.97 million €). Firm profits also rise by 1.41 million €, and tax revenue increases by 1.48 million €. These results indicate

that the two-threshold design outperforms both the one-threshold and linear designs in terms of total welfare, consumer benefits, and externality reduction.

Table 3.1: Welfare variation decomposition

Counterfactual	Change (millions euros) in				
	Total welfare ΔW_t	Firm profit ΔW_t^π	Consumer surplus ΔW_t^C	Tax revenue ΔW_t^{tax}	Externality ΔW_t^{SC}
<i>Linear design</i> 27€\kg of sugar	5.64	-0.45	-6.45	4.17	8.37
<i>Tiered design</i>					
1 threshold	10.11	0.41	3.06	0.00	6.64
2 thresholds	21.44	1.41	7.97	1.48	10.57
SDIL	5.63	-0.49	-4.64	2.82	7.94

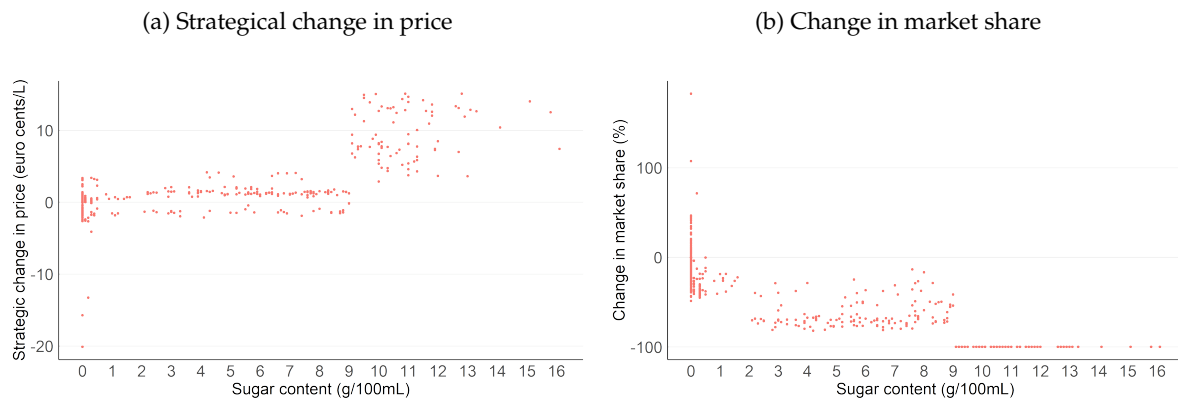
Notes: This table plots the changes in total welfare (column 2) and in the different components of total welfare, assuming that firms have price strategic behavior: firms' profit (column 3), consumer surplus (column 4), tax revenue (column 5) and externality (column 6). The tiered designs with 1 and 2 thresholds are empirically derived from the procedure detailed in section 3.3.1 when we include firms' strategic behavior.

Mechanisms. Figure 3.2a illustrates the strategic price adjustments made by firms following the implementation of the two-threshold optimal design. For products with sugar content exceeding 9 grams per 100 mL, firms choose to increase prices by more than the imposed tax. This pricing behavior results in a nearly 100% loss in market share for these high-sugar products (Figure 3.2b). In this context, the optimal design operates as a corner solution, where the tax at the second threshold is set high enough to mostly drive high-sugar products out of the market.

For other products, we do not observe any clear differences between those with 2 to 9 grams of sugar per 100 mL, which are taxed at 15 cents, and those with less than 2 grams, which are exempt from the tax. Interestingly, some products without sugar see an important decrease in price.

By lowering prices on low-sugar products, firms make these healthier options more accessible to consumers, thereby improving consumer surplus. Importantly, this strategic price reduction does not come at the expense of firm profits for these products. Rather, as shown in Figures 3.2a and 3.2b, firms' adjustments increase the market share of low-sugar products.

Figure 3.2: Effects on prices and market shares



Notes: These figures plot the change in price and market share along the sugar content distribution following the implementation of the tax. The y-axis plots the price change (in euro cents per L, after removing the tax amount) in Figure (a) and the market share change (in %) in Figure (b). The x-axis displays the sugar content of products (in grams per 100 mL). The difference in price is deflated from the value of the tax, in order to highlight how firms strategically adapts.

4.2 Validity of our methodology

We propose a design that outperforms the linear design, which is theoretically considered the best, as it taxes each product according to its marginal externality. This surprising result arises from the fact that the theoretical optimal linear design does not account for the competitive environment, which leads to a new equilibrium. In the next section, we will study the role of incorporating strategic pricing into the regulator's maximization program. We will also emphasize the importance of carefully estimating the social cost of sugar in designing taxes.

4.2.1 Comparison of strategic vs. non-strategic tax designs

The comparison between tax designs with and without firms' strategic responses illustrates the critical importance of accounting for firm behavior in achieving optimal welfare outcomes.

Description of the design. We derive the optimal designs for one-tiered and two-tiered taxes under the assumption that firms fully pass on the tax to consumers. In the one-threshold design, only products with sugar content above 4 grams per 100 mL are taxed at a flat rate of 0.21 €. In the two-threshold design, products with less than 2 grams of sugar per 100 mL are exempt from the tax, products with sugar content between 2 and 7 grams per 100 mL are taxed at 0.12 €, and products with more than 7 grams of sugar per 100 mL are taxed at 0.26 € (Figure 3.D.1). The tax levels are

significantly lower and more in line with those implemented in practice when strategic firm behavior is not taken into account.

Welfare decomposition. In the absence of strategic pricing responses, where firms fully pass the tax onto consumers, the welfare gains from taxation are notably smaller (Table 3.2). Without strategic adjustments, the tax is less effective in changing consumption patterns, as consumers are not sufficiently incentivized to shift towards lower-sugar products in response to price changes. It is also important to note that without firms' strategic responses, the linear design performs better than the optimal tiered designs with one or two thresholds.

In contrast, when firms react strategically by adjusting their prices, several important dynamics arise. First, consumer surplus increases as firms reduce prices on low-sugar products, which would not occur in a non-strategic environment. This price reduction makes healthier products more attractive to consumers, thereby enhancing welfare in a way that would be unattainable under non-strategic assumptions. Second, despite the tax, firms can maintain or even improve profits by decreasing the prices of low-sugar products, compensating for any potential losses in high-sugar product sales. Firms that strategically adjust their pricing are able to sustain profits through increased demand for lower-sugar alternatives. Thus, incorporating strategic firm behavior into the design of sugar taxes is key to finding optimal design.

Table 3.2: Welfare variation decomposition without firms' strategic behavior

Counterfactual	Change (millions euros) in				
	Total welfare ΔW_t	Firm profit ΔW_t^π	Consumer surplus ΔW_t^C	Tax revenue ΔW_t^{tax}	Externality ΔW_t^{SC}
<i>Linear design</i> 27€\kg of sugar	4.51	-0.25	-7.79	4.37	8.17
<i>Tiered design</i>					
1 threshold	4.11	-0.25	-6.69	3.59	7.46
2 thresholds	4.37	-0.26	-7.08	3.91	7.80
SDIL	4.10	-0.16	-6.36	3.00	7.63

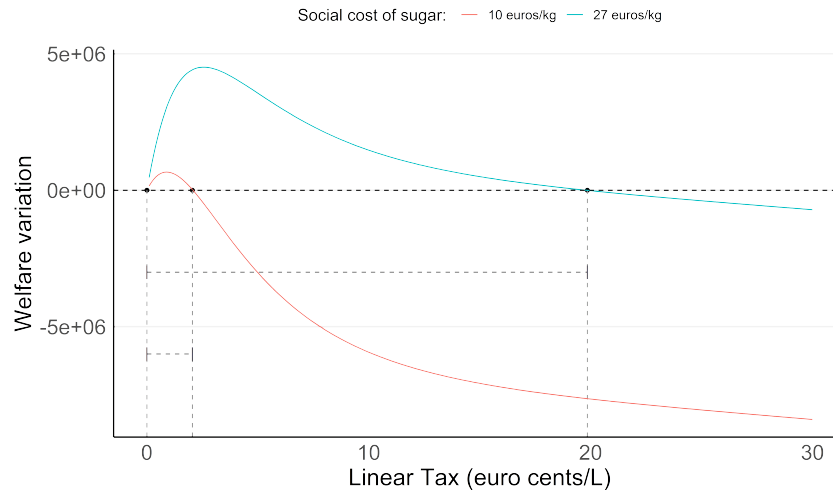
Notes: This table plots the changes in total welfare (column 2) and in the different components of total welfare, assuming no firms' price strategic behavior: firms' profit (column 3), consumer surplus (column 4), tax revenue (column 5) and externality (column 6). The tiered designs with 1 and 2 thresholds are empirically derived from the procedure detailed in section 3.3.1 when we disregard firms' strategic behavior.

4.2.2 Discussion on the value of the social cost of sugar

In this subsection, we examine how our calibration of social costs influences the results. We consider two levels of social costs: 10 euros per kg (lower bound) and 27 euros per kg (average).²²

We begin by comparing the welfare variations associated with the two estimates in a linear design. As the social cost of sugar rises from 10 to 27 € per kg, the highest welfare gain also increases from nearly 0.7 to a bit more than 4.5 million € (Figure 3.3). Additionally, we observe that the range of tax rates resulting in welfare improvements significantly expands with an increase in the social cost of sugar. While only tax rates below 2.1 € cents lead to welfare improvements with a social cost of 10 € per kg, this range extends to tax rates up to 20 € cents with a social cost of 27 € per kg. Therefore, the magnitude of the externality cost plays a critical role in determining the optimal design and the potential necessity for intervention.

Figure 3.3: Welfare change in linear taxation with tax rate



Notes: This figure plots the variation in welfare that would result from implementing a linear tax according to different tax rates for two values of the social cost of sugar. The y-axis plots the variation in welfare (in millions of euros) and the x-axis displays the different tax rates for a linear tax (in euros cents per liter). The red line corresponds to the social cost estimated at 10 euros per kg and the blue line corresponds to the social cost estimated at 27 euros per kg.

4.3 Comparison with the SDIL

The Soft Drinks Industry Levy. In 2018, the United Kingdom took a significant step in addressing the public health concerns surrounding excessive sugar consumption by

²²We disregard here firms' strategic behaviors because we focus on the tax implementation and not the optimal designs.

implementing the Soft Drinks Industry Levy on sugar-sweetened beverages. This tax was part of the government's broader strategy to tackle rising rates of obesity and related health issues, particularly among children. The design of the tax is characterized by a two-tiered approach based on the total sugar content of products. Beverages with a sugar content above 8 grams of sugar per 100 mL face a high tax rate (24 pounds per liter), while those with sugar content between 5 to 8 grams per 100 mL are subject to a lower rate (18 pounds per liter) and those with less than 5 grams of sugar per 100 mL are exempt from the tax (Figure 3.A.4).²³ Pure fruit juices and drinks with high milk content are exempt from the tax.

Assessment. We now assess the welfare effects of the implemented SDIL. The SDIL, when evaluated under the assumption that firms do not engage in strategic pricing, results in a welfare gain of 4.10 million €, a result that is close to the 4.51 million € gain associated with a linear tax design (Table 3.2). However, when strategic responses from firms are considered, the welfare performance of the SDIL improves slightly, closing the gap with the linear tax design. The SDIL achieves a welfare gain of 5.63 million €, while the linear tax results in a gain of 5.64 million € (Table 3.1).

Although the SDIL performs reasonably well compared to the linear tax, it lags behind the welfare gains achieved by the two optimal tax designs in the presence of strategic pricing. The SDIL's structure, while effective at raising tax revenue and moderately shifting consumption patterns, could be improved by better targeting higher-sugar products. These results suggest that while the SDIL provides a good baseline for sugar taxation, more optimal designs can be achieved by incorporating strategic pricing behavior to reach better welfare outcomes. Indeed, the choice of threshold locations has little impact on welfare when not considering firms' price reactions, unlike in the more realistic case where firms respond to prices (Figure 3.D.2).

²³The following conversion rate is used: 1 pound = 1.15 euros.

5 Conclusion

Many sugar taxes have been implemented in the form of tiered taxes, contradicting what is advocated by theoretical and empirical literature. In this paper, we investigate how tiered tax designs would perform compared to the theoretical optimal linear tax. We introduce a comprehensive framework for evaluating the effect of tax designs on total welfare, accounting for the externality generated by excess sugar consumption, the heterogeneity in patterns of consumption and firms' strategic behavior. Our analysis shows that optimal sugar tax designs which account for firms' strategic pricing lead to significantly greater welfare improvements than non-strategic designs or the implemented SDIL. The optimal tax design, by imposing higher taxes on high-sugar products and creating a new equilibrium where firms reduce prices on low-sugar products, results in both better public health outcomes and higher consumer surplus. Additionally, firms are able to preserve profits through strategic price adjustments. Policymakers seeking to maximize the welfare impact of sugar taxes should, therefore, carefully consider firms' strategic responses when designing their taxation policies.

In future work, we aim to extend this analysis in several ways. First, we will explore the heterogeneous effects of sugar taxes across demographics, allowing for a more nuanced understanding of the policy's impact. We also plan to investigate the heterogeneous effects across firms to better capture industry-level responses to taxation. Additionally, we will replicate this analysis in other European countries, such as France and Spain, which have also implemented tiered taxes on soft drinks. Lastly, we intend to assess whether our findings can be generalized to other markets, such as the biscuit industry, which remains untaxed despite its considerable contribution to sugar consumption.

Appendices to Chapter 3

3.A Data

3.A.1 Demographics

Definition of obese and overweight. Adults with BMI, defined as weight (in kg) divided by the square of the height (in m), above 30 are considered as obese, and individuals with BMI ranging between 25 and 30 are considered as overweight. The final dataset contains only households with no missing value on adults' BMI in the analysis. We exclude 6,353 households (21%).

Definition of socio-economic class. It is based on the socio-professional categories.

Table 3.A.1: Household characteristics in Kantar panel

	N	%
Household composition		
Without children	16,093	65
With children below 6 years old	2,960	12
With children 7-16 years old	3,668	15
With children both below 6 and 7-16 years old	1,865	8
Obesity status		
No overweight or obese adults	5,972	24
Some overweight or obese adults	8,184	33
All overweight or obese adults	10,430	43
Socio-economic class		
Rich	5,262	21
Average	13,912	57
Poor	5,412	22
All	24,586	

Notes: This table shows the characteristics of households in the panel according to their family composition, obesity status and socio-economic class.

3.A.2 Market definition

Table 3.A.2: Summary of the market definition

Sub-categories	Included?
Sugar-sweetened beverage (<i>cola, iced tea, lemonade, fruit-flavored drink, sport and energy drink, tonic, other</i>)	yes
Fruit juice (<i>pure fruit juice, nectar, smoothie, fruit juice with milk</i>)	yes
Flavoured water	yes
Milk-based drink	yes
Water	no
Syrup, powdered drink, cocktail mixer	no

Notes: This table summarizes which products are included in our analysis.

3.A.3 Products

3.A.3.1 Details on the definition of products

Sub-categories. The sugar-sweetened beverage category is made up of colas, iced tea, fruit-flavored drinks, flavored water, tonic water, lemonade, energy drinks, and other SSBs. The fruit juice category includes nectar, fruit juice made from concentrate, pure fruit juices and smoothies. Milk-based drinks are made up of flavored milk.

Added sugar. The dataset flags products containing added sugar.

Additional details on the product definition. Beverages of a given sub-category with a small purchase occurrence are aggregated either with similar products of the same firm or in a hypothetical product defined as other firm/other brand of the sub-category. Additionally, private labels are aggregated into one firm and one brand.

3.A.3.2 Market structure

Table 3.A.3 presents a summary of the market structure with the definition of alternatives. The UK market is characterized by a high number of national brands, resulting in a high number of non-alcoholic beverage alternatives. Mean prices in Table 3.A.3 are calculated as the weighted ratio of total expenditure in euros over the total quantities in liters of all drinks purchased.

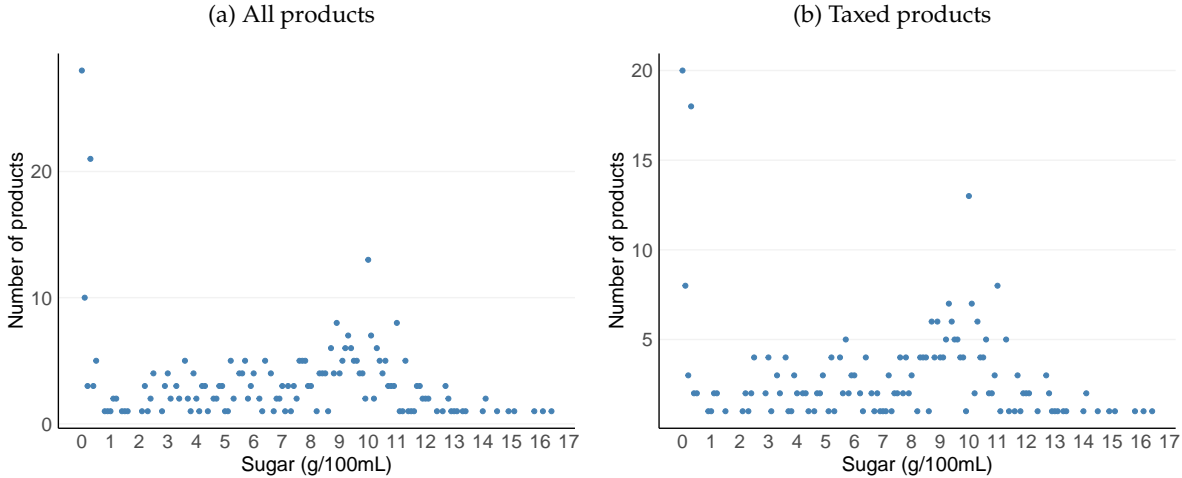
Table 3.A.3: Market structure

Number of firms	78
Number of national brands	151
Number of sub-categories	11
Number of alternatives	402
Number of observations	1,319,069
Mean price (per liter)	0.78 €
Mean price of non-alcoholic beverages with added sugar (per liter)	0.67 €

Notes: This table provides some descriptive statistics about the market structure in our analysis. The number of observations refers to the number of Kantar references purchased per household, date and store. If on a given date and in a given store, a household buys 3 bottles of 1L of Coke, this is an observation, but if a household buys 2 bottles of 1L of Coke and 1 bottle of 1.5L, this is 2 observations.

3.A.3.3 Distribution of the sugar content

Figure 3.A.1: Distribution of the sugar content of the non-alcoholic beverage supply



Notes: These figures show the distribution of sugar content for all products and for only taxed products (supply dimension). Specifically, they display the number of products every 0.1 grams of sugar per 100mL.

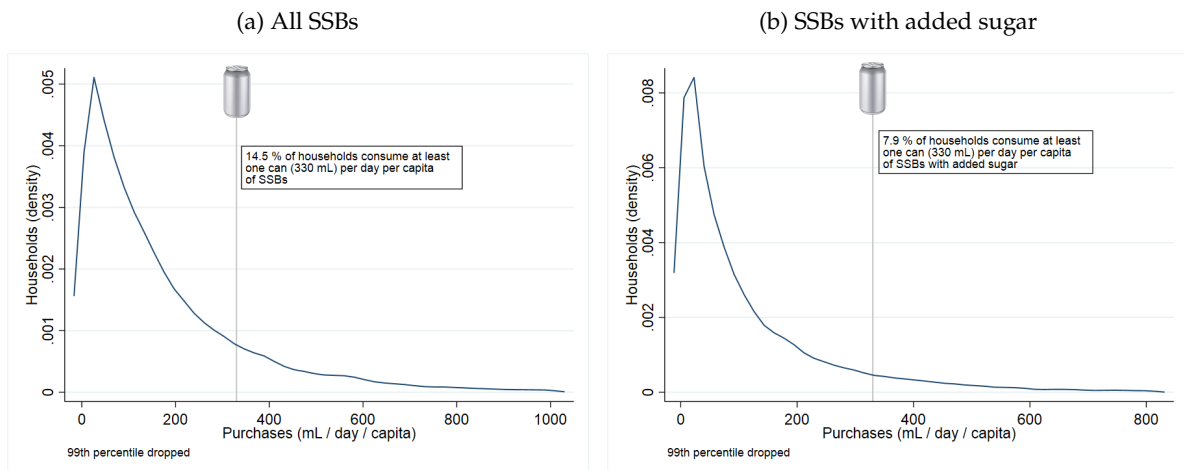
3.A.3.4 Descriptive statistics

Table 3.A.4: Average market share, price and sugar content of each sub-category

	Share %	Price €/l	Sugar g/100ml
<i>Sugar-sweetened beverages</i>			
Colas	26.8	0.8 (1.0)	3.1 (5.0)
Regular	29.6	0.8 (1.3)	10.3 (1.5)
Diet	70.4	0.8 (0.4)	0.0 (0.0)
Fruit-flavoured drinks	12.2	0.9 (0.8)	3.4 (3.8)
Regular	73.2	1.2 (0.6)	6.1 (3.2)
Diet	26.8	0.5 (1.1)	0.4 (2.3)
Flavoured waters	22.2	0.4 (0.7)	0.3 (1.3)
Regular	1.4	0.8 (1.1)	4.7 (1.6)
Diet	98.6	0.4 (0.6)	0.2 (0.1)
Tonic waters	3.1	0.9 (1.2)	1.8 (2.7)
Regular	45.8	1.2 (1.3)	4.3 (1.8)
Diet	54.2	0.7 (1.3)	0.3 (1.4)
Lemonades	5.0	0.4 (1.1)	1.8 (4.2)
Regular	40.0	0.5 (1.3)	4.9 (3.3)
Diet	60.0	0.3 (0.3)	0.1 (0.1)
Energy drinks	5.8	1.4 (0.9)	5.4 (5.2)
Regular	74.5	1.4 (0.9)	7.0 (3.1)
Diet	25.5	1.5 (1.0)	0.2 (0.2)
Other SSBs	3.5	0.7 (1.1)	3.6 (4.4)
Regular	46.7	1.0 (1.1)	9.4 (2.5)
Diet	53.3	0.6 (1.2)	0.1 (1.6)
<i>Fruit juices</i>			
Nectars	5.6	1.1 (0.6)	7.2 (3.5)
Regular	70.9	1.1 (0.6)	7.2 (3.5)
Diet (no added sugar)	29.1	0.9 (0.4)	2.5 (2.2)
Pure fruit juice	11.7	1.2 (0.9)	9.8 (2.7)
Smoothies	1.5	2.9 (0.9)	10.7 (1.2)
Regular	72.7	2.7 (0.7)	10.5 (1.1)
Diet (no added sugar)	27.3	3.6 (0.9)	11.2 (1.5)
<i>Milk-based drinks</i>			
Flavoured milks	2.6	1.7 (1.3)	9.2 (2.5)
Diet	92.9	1.7 (1.4)	9.4 (2.3)
Regular	7.1	2.0 (1.1)	5.6 (2.5)

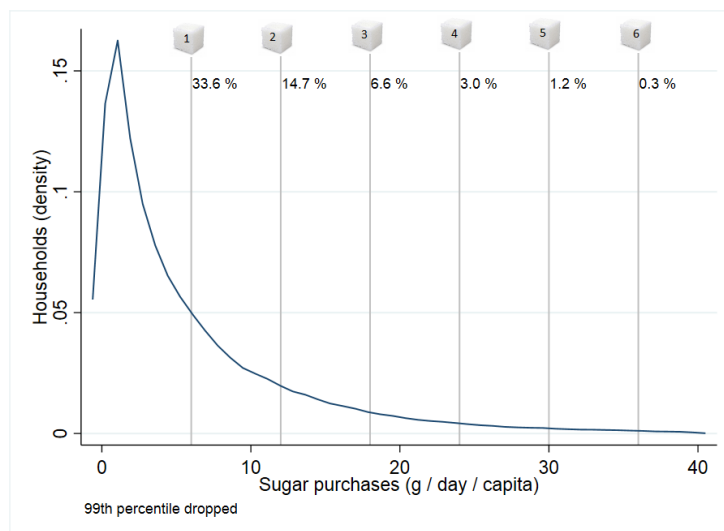
Notes: This table provides descriptive statistics for the different sub-categories or products. Column 2 displays the market share, column 3 the average price (in euros per liter) and column 4 the average sugar content (in grams per 100mL). Standard deviation are in parentheses.

Figure 3.A.2: Purchase of non-alcoholic beverages across Kantar households



Notes: These figures show the distribution of households according to their purchases of non-alcoholic beverages (in mL per day per capita). Figure (a) is made on the purchases of all beverages and Figure (b) only on beverages containing added sugar.

Figure 3.A.3: Sugar purchases across Kantar households



Notes: This figure shows the distribution of households according to the sugar content of their purchases of non-alcoholic beverages (in grams per day per capita). 33.6 % of households purchase more than one sugar cube per day per capita.

Table 3.A.5 provides a comparison of the purchase of non-alcoholic beverages per capita per year across households with respect to household composition, obesity status, and socio-economic class. Households with children between 7 and 16 years old are the highest consumers and households with children between 0 and 6 years old are the lowest consumers. The mean purchase increases gradually with the proportion of adults who are overweight or obese within the household. Furthermore, poor households buy more than rich households.

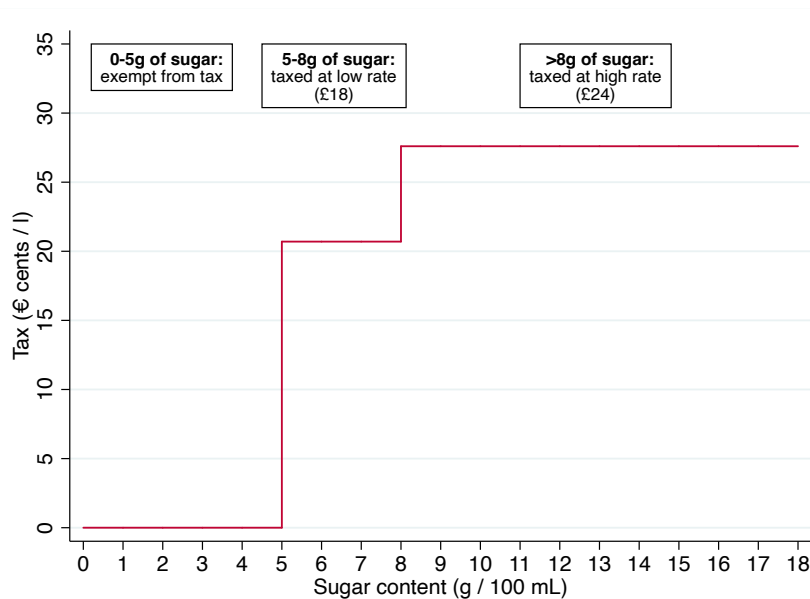
Table 3.A.5: Household purchases

Mean quantity (l/capita/year)	All beverages	Added sugar
Household composition		
No children	67 (85)	46 (68)
Children below 6	45 (54)	30 (44)
Children 7-16	76 (74)	53 (60)
Children both below 6 & 7-16	53 (57)	36 (46)
Obesity status		
None overweight or obese	53 (73)	37 (61)
Some overweight or obese	63 (71)	44 (58)
All overweight or obese	73 (87)	50 (68)
Socio-economic class		
Rich	59 (68)	38 (55)
Average	65 (79)	45 (62)
Poor	70 (88)	51 (73)
All	65 (79)	44 (64)

Notes: This table shows the purchases of non-alcoholic beverages (in liters per capita per year) for the different types of households. Column 2 displays the quantities for all beverages and column 3 for products containing added sugar. Standard deviations are in parentheses.

3.A.4 Institutional background (SDIL)

Figure 3.A.4: Design of the SDIL



Notes: This figure describes the design of the SDIL tax implemented in the UK. The y-axis plots the tax rate (in euro cents per L) and the x-axis represents the sugar content of products (in grams per 100 mL).

Tax implementation. Manufacturers are responsible for measuring and reporting the sugar content of their products to the tax authorities. The tax authorities then apply the relevant tax rate based on the sugar content reported by the manufacturer. Regulatory agencies may perform audits or random checks to verify the accuracy of the reported sugar content. If discrepancies are found (for example, if a manufacturer underreports

the sugar content to avoid higher tax rates), penalties may be imposed.

3.B Demand

3.B.1 Price equation

Table 3.B.1: Results on price equation

	Coefficient (se)
Instrumental variables	
<i>BLP instruments</i>	
Number of competing products offered by other firms within the product category	-0.01*** (0.00)
Total sugar content of competing products within the nutritional category	-0.00*** (0.00)
<i>Cost shifter (input prices)</i>	
Glass	-0.00*** (0.00)
Exogenous variables	
Diet	0.11*** (0.04)
Sugar (Soft drinks)	0.07*** (0.01)
Sugar (Fruit juices)	0.28*** (0.01)
Brand fixed effects	yes
Category fixed effects	yes
IV joint significance test	F(3,5035) = 35.97 Prob > F = 0.0000
Observations	5,182
R ²	0.944

Notes: This table displays the coefficients of the price equation where we regress prices on instrumental variables and exogenous variables of the demand equation. Estimators' standard errors (se) are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.B.2 Estimation

We estimate the demand model using individual data. We use the simulated maximum likelihood method as in [Revelt and Train \[1998\]](#). In our model, as purchases are independent considering household and period, the likelihood function for purchase i can be written as:

$$L_i = \int \prod_{t=1}^T \prod_{j=0}^{J_t} \left[\frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J_t} \exp(\delta_{kt} + \mu_{ikt})} \right]^{Y_{ijt}} dP_v(v) \quad (3.B.1)$$

where Y_{ijt} is a dummy variable equal to 1 if purchase i involves product j in period t and 0 otherwise.

We estimate the vector of demand parameters by maximizing the simulated log-likelihood

function given by:

$$SLL = \sum_{i=1}^N W_i \times \ln[\hat{s}_{ijt}] \quad (3.B.2)$$

where N is the total number of purchases registered by Kantar for a given market and $W_i = N_i \times W_i^{ht}$ is the weight associated to the purchase i . N_i is the number of units purchased at the i^{th} purchase, used to take into account multiple choices of the same type of product (*i.e.*, buying three bottles of the same product at the same time). This is then multiplied by Kantar period-specific household sample weights associated to the household h that made the i^{th} purchase W_i^{ht} .

\hat{s}_{ijt} stands for the individual simulated market share of product j in period t :

$$\hat{s}_{ijt} = \frac{1}{R} \sum_{r=1}^R w_r \prod_{t=1}^T \prod_{j=0}^{J_t} \left[\frac{\exp(\alpha_i^r p_{jt} + \beta_i X_{jt} + v \hat{\eta}_{jt})}{\sum_{k=1}^{J_t} 1 + \exp(\alpha_i^r p_{kt} + \beta_i X_{kt} + v \hat{\eta}_{kt})} \right]^{Y_{ijt}} \quad (3.B.3)$$

where R is the number of draws, w_r is the r^{th} entry in the vector of weights and $\alpha_i^r = \alpha + \pi^\alpha D_i + \sigma v_i^r$ with $v^r \sim \mathcal{N}(0, 1)$.²⁴ $\hat{\eta}_{jt}$ is the estimated error term of product j in period t from the price equation (Table 3.B.1).

3.B.3 Results

We found that households with young children are less sensitive to price and households with children above 7 years of age are more sensitive to price than households without children. Households with overweight or obese individuals are more sensitive to price than households with no overweight or obese individuals. Results also suggest that households prefer diet to regular products. The preference for diet products is stronger for households with children. Households with children have a higher taste for sugar. For a given brand and a given regular or diet characteristic, households prefer products with more sugar both for SSBs and fruit juices.²⁵

²⁴To address the curse of dimensionality, we use a sparse grid method as it is developed in Heiss and Winschel [2008]. The choice for the number of simulations is explained on <http://www.sparse-grids.de/>. Integration on sparse grids has low computational costs compared to other methods. The number of draws is especially reduced because each draw is associated with a weight. Let R be the number of simulations which depends on the type of integration rule, the number of dimensions and the chosen accuracy. Let x_r and w_r be the r^{th} entries respectively in the vector of draws and the vector of weights. We want v^r to follow a standard normal distribution so we apply a transformation (inverse of the standard normal cumulative distribution function) to x_r ($x_r \in [0, 1] \forall r$) so that $v^r = \Phi^{-1}(x_r)$ and $v^r \sim \mathcal{N}(0, 1)$.

²⁵The sugar coefficients do not represent the absolute preference for sugar but rather the preference between products for a given brand. Indeed, when controlling for brand fixed effects, the remaining variation in sugar content comes from the different products sold by the same brand (especially products from different categories, *e.g.*, "fruit-flavored drinks" and "other SSBs").

Table 3.B.2: Estimates of the random coefficient logit model

Price (p_{jt})	
Mean (α)	-6.84 (0.00)
× children below 6 years old	0.13 (0.00)
× children 7-16 years old	-0.03 (0.00)
× average class	-0.24 (0.00)
× poor class	-0.35 (0.00)
× at least one obese	-0.12 (0.00)
× all obese	-0.15 (0.00)
Standard deviation (σ)	1.64 (0.00)
Diet	
× children below 6 years old	0.13 (0.00)
× children 7-16 years old	0.07 (0.00)
Sugar (SSBs)	
× children below 6 years old	0.03 (0.00)
× children 7-16 years old	0.02 (0.00)
Sugar (fruit juices)	
× children below 6 years old	0.05 (0.00)
× children 7-16 years old	0.04 (0.00)
Fixed effects	
Sub-category	yes
Brand (NBs & PLs)	yes
Error ($\hat{\eta}_{jt}$)	4.09 (0.00)
Observations: 1,364,000	
Log-likelihood: -2.32081e+10	

Notes: This table displays the results from the demand estimation detailed in section 3.2.1. Standard errors of coefficient estimates are in parentheses.

Table 3.B.3: Own-price elasticities by household characteristics

	Own-price elasticities
Household composition	
Without children	-8.93
With children below 6 years old	-8.84
With children 7-16 years old	-9.01
With children below 6 and 7-16 years old	-8.92
Obesity status	
No overweight or obese	-8.84
At least one overweight or obese	-8.95
All overweight or obese	-8.98
Socio-economic class	
Rich	-8.73
Average	-8.96
Poor	-9.08

Notes: This table displays the average own-price elasticities by household characteristics. Own-price elasticities are calculated at the alternative level. For a given household characteristic, it is the mean elasticity computed across all the alternatives' own-price elasticities.

Table 3.B.4: Own-price elasticities by sugar content and household characteristics

	Sugar content (g/100mL)					
	0] 0 ; 5 [[5 ; 8 [[8 ; 10 [[10 ; 12 [≥ 12
All households	-7.17	-6.90	-9.53	-10.18	-10.13	-10.12
<i>Household composition</i>						
Without children	-7.16	-6.89	-9.54	-10.19	-10.14	-10.13
With children below 6	-7.11	-6.84	-9.44	-10.08	-10.03	-10.03
With children 7-16	-7.23	-6.95	-9.62	-10.27	-10.22	-10.21
With children below 6 & 7-16	-7.18	-6.90	-9.52	-10.16	-10.11	-10.11
<i>Obesity status</i>						
No overweight or obese	-7.11	-6.84	-9.44	-10.07	-10.03	-10.03
At least one overweight or obese	-7.19	-6.91	-9.56	-10.21	-10.16	-10.15
All overweight or obese	-7.22	-6.94	-9.60	-10.24	-10.19	-10.19
<i>Socio-economic class</i>						
Rich	-7.03	-6.76	-9.32	-9.95	-9.90	-9.91
Average	-7.20	-6.92	-9.57	-10.22	-10.17	-10.16
Poor	-7.29	-7.01	-9.71	-10.36	-10.31	-10.30

Notes: This table displays the own-price elasticities for different brackets of sugar content according to household characteristics.

3.C Supply

Table 3.C.1: Elasticities and margins by firms

Firms	Own-price elasticities	Margins (% price)
National brands		
Firm 1	-6.5	19.2
Firm 2	-6.6	17.6
Firm 3	-8.2	12.6
Firm 4	-4.2	27.2
Firm 5	-5.6	19.5
Firm 6	-3.5	28.4
Firm 7	-4.8	21.4
Firm 8	-12.5	8.7
Firm 9	-10.8	9.8
Firm 10	-6.4	16.5
Small firms	-7.9	18.9
Very small firms	-7.5	15.3
Private labels	-4.4	52.3

Notes: This table displays the own-price elasticities and the margins (as a percentage of price) for national brands and private labels. Firm 1 to Firm 10 are the ten firms with the highest market shares.

Table 3.C.2: Margins by sugar content

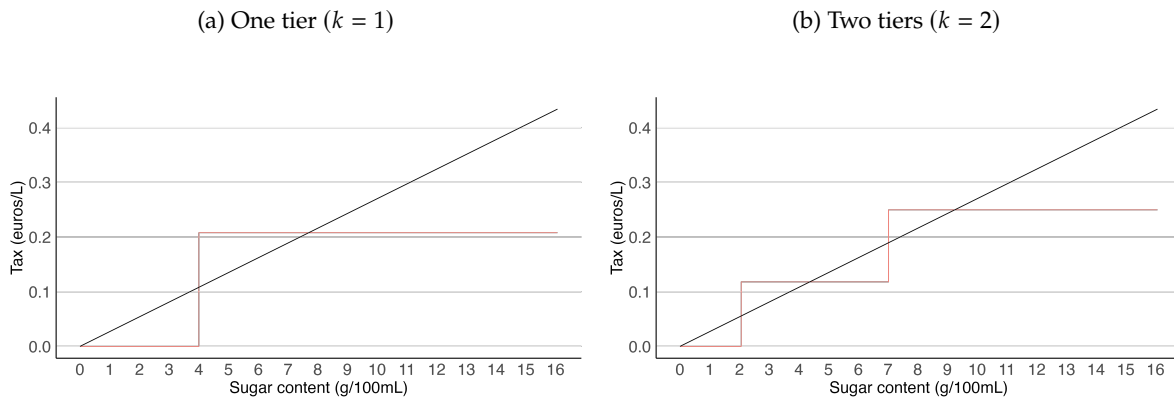
Sugar content (g/100mL)	Margins (% price)
0	19.2
] 0 ; 5 [49.9
[5 ; 8 [20.1
[8 ; 10 [17.7
[10 ; 12 [22.5
≥ 12	19.1

Notes: This table displays the margins (as a percentage of price) for different brackets of sugar content.

3.D Counterfactual simulations

3.D.1 Optimal designs without firms' strategic behavior

Figure 3.D.1: Comparison of tiered tax designs (without firms' strategic behavior)



Notes: The two figures plot the optimal tiered tax design with one and two thresholds respectively (in red) where we ignore firms' strategic behavior. The optimal tax rates and location of thresholds are obtained empirically from the maximization program presented in section 3.3.1. The black line corresponds to the theoretical optimal linear tax design where the tax rate is set at the value of the social cost (27 euros per kg of sugar). The y-axis plots the tax rate (in euros cents per L) and the x-axis displays the sugar content of products (in grams per 100 mL).

3.D.2 Optimization

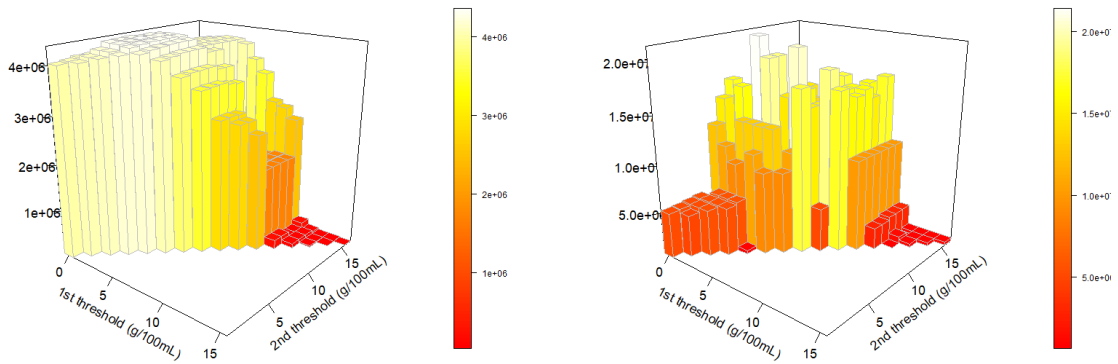
The regulator's choice of thresholds becomes crucial when firms' strategic pricing behavior is considered. Figure 3.D.2 plots the maximum achievable welfare across different threshold locations for two-tiered designs.²⁶ When considering non-strategic firms, welfare variation remains relatively stable around the optimal threshold (Figure 3.D.2b). However, when considering firms' strategic responses, the variations become more pronounced (Figure 3.D.2a).

²⁶See Figures 3.D.3 and 3.D.4 for similar graphs with one threshold and with two thresholds in 2D.

Figure 3.D.2: Welfare variation according to the location of thresholds

(a) No strategic pricing

(b) Strategic pricing

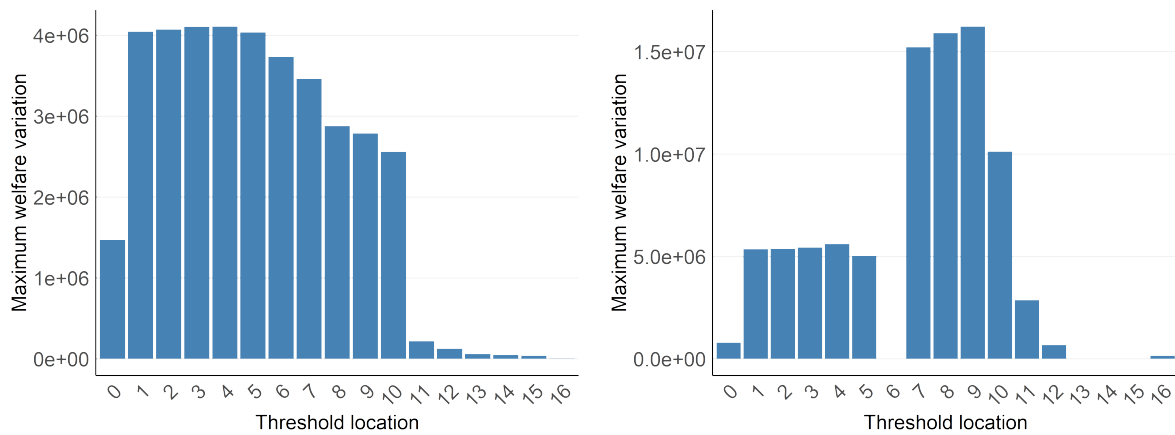


Notes: These figures plot the variation in welfare (in euros) according to the location of the two thresholds (in grams per 100 mL). The lighter the colors, the higher the welfare gains. Figure (a) disregards firms' strategic behavior and Figure (b) accounts for it.

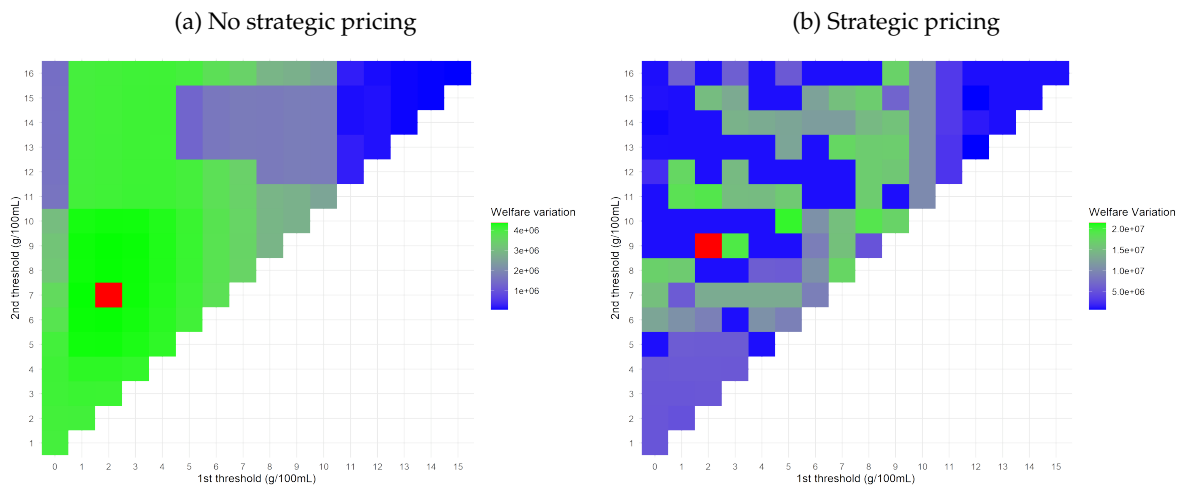
Figure 3.D.3: Welfare variation according to the location of threshold ($k = 1$)

(a) No strategic pricing

(b) Strategic pricing



Notes: These figures plot the variation in welfare (in euros) according to the location of the threshold (in grams per 100 mL). Figure (a) disregards firms' strategic behavior and Figure (b) accounts for it.

Figure 3.D.4: Welfare variation according to the location of thresholds ($k = 2$)

Notes: These figures plot the variation in welfare (in euros) according to the location of the two thresholds (in grams per 100 mL). The red square corresponds to the threshold location that achieves the highest welfare gains. Figure (a) disregards firms' strategic behavior and Figure (b) accounts for it.

Bibliography

Abeliansky, Ana Lucia, Matthias Beulmann, and Klaus Prettnner, “Are they coming for us? Industrial robots and the mental health of workers,” *Research Policy*, April 2024, 53 (3), 104956.

Acemoglu, Daron and David Autor, “Tasks and Technologies: Implications for Employment and Earnings,” *Handbook of Labor Economics*, 2011, 4, 1043–1171.

– **and Pascual Restrepo**, “The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment,” *American Economic Review*, 2018, 108, 1488–1542.

– **and –**, “Artificial Intelligence, Automation, and Work,” *The Economics of Artificial Intelligence: An Agenda*, 2019, pp. 197–236.

– **and –**, “Robots and Jobs: Evidence from US Labor Markets,” *Journal of Political Economy*, 2020, 128 (6), 2553–2584.

– **and –**, “Demographics and Automation,” *The Review of Economic Studies*, 1 2022, 89, 1–44.

– **, Claire Lelarge, and Pascual Restrepo**, “Competing with Robots: Firm-Level Evidence from France,” *AEA Papers and Proceedings*, 5 2020, 110, 383–388.

Aghion, Philippe, Céline Antonin, Simon Bunel, and Xavier Jaravel, “The Local Labor Market Effects of Modern Manufacturing Capital: Evidence from France,” *American Economic Association: Papers and Proceedings*, 2023, 113, 219–223.

– , – , – , **and –**, “Modern Manufacturing Capital, Labor Demand, and Product Market

-
- Dynamics: Evidence from France," *Working paper*, 2024.
- Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky**, "Regressive Sin Taxes, with an Application to the Optimal Soda Tax," *The Quarterly Journal of Economics*, 05 2019, 134 (3), 1557–1626.
- , –, and –, "Should We Tax Sugar-Sweetened Beverages? An Overview of Theory and Evidence," *Journal of Economic Perspectives*, 2019, 33 (3), 202–227.
- Arntz, Melanie, Sebastian Findeisen, Stephan Maurer, and Oliver Schlenker**, "Are We Yet Sick of New Technologies? The Unequal Health Effects of Digitalization," *Working paper*, 2024.
- Athey, Susan, Lisa K. Simon, Oskar N. Skans, Johan Vikstrom, and Yaroslav Yakymovych**, "The Heterogeneous Earnings Impact of Job Loss Across Workers, Establishments, and Markets," *IFAU Working Paper*, 2024, 10.
- Autor, David**, "Why are there still so many jobs? the history and future of workplace automation," *Journal of Economic Perspectives*, 2015, 29, 3–30.
- and **David Dorn**, "The Growth of Low-Skill Service Jobs and the Polarization of the U.S. Labor Market," *American Economic Review*, 2013, 103, 1553–1597.
- , **Frank Levy, and Richard Murnane**, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, 118, 1279–1333.
- Barlet, Muriel, Magali Coldefy, Clémentine Collin, and Véronique Lucas-Gabrielli**, "Local Potential Accessibility (LPA): A New Measure of Accessibility to Private General Practitioners," *Questions d'Economie de la Santé*, 2012, 174.
- Behrens, Kristian and Julien Martin**, "Concording large datasets over time: The C3 method," 2015. Technical document (Unpublished).
- Bergemann, Anne, Erik Grönqvist, and Sofia Guobjörnsdóttir**, "Diabetes Morbidity After Displacement," *Health and Labor Markets (Research in Labor Economics)*, 2019, 47, 99–154.
- Bergounhon, Flora, Clémence Lenoir, and Isabelle Méjean**, "A guideline to French firm-level trade data," 2018. Technical document (Unpublished).

-
- Berry, Steven**, “Estimating discrete-choice models of product differentiation,” *RAND Journal of Economics*, 1994, 25, 242–262.
- , **James Levinsohn**, and **Ariel Pakes**, “Automobile Prices in Market Equilibrium,” *Econometrica*, 1995, 63 (4), 841–890.
- , – , and – , “Differentiated products demand systems from a combination of micro and macro data: The new car market,” *Journal of political Economy*, 2004, 112 (1), 68–105.
- Bertheau, Antoine**, **Edoardo Acabbi**, **Cristina Barceló**, **Andreas Gulyas**, **Stefano Lombardi**, and **Raffaele Saggio**, “The Unequal Cost of Job Loss across Countries,” *American Economic Review: Insights (Forthcoming)*, 2022.
- Bertoni, Marco**, **Giorgio Brunello**, and **Filippo. Da Re.**, “Pension reforms, longer working horizons and depression. Does the risk of automation matter?,” *Labour Economics*, 2023, 85.
- Bessen, James**, **Maarten Goos**, **Anna Salomons**, and **Wiljan van den Berge**, “What Happens to Workers at Firms that Automate?,” *The Review of Economics and Statistics (forthcoming)*, 2023.
- Bhattacharya, Jay** and **Neeraj Sood**, “Who Pays for Obesity?,” *Journal of Economic Perspectives*, 2011, 2 (25), 139–158.
- Blasco, Sylvie**, **Julie Rochut**, and **Bénédicte Rouland**, “Displaced or Depressed? The Effect of Working in Automatable Jobs on Mental Health,” *IZA Discussion Paper Series*, 2022.
- Bloemen, Hans**, **Stefan Hochguertel**, and **Jochem Zweerink**, “Job loss, firm-level heterogeneity and mortality: Evidence from administrative data,” *Journal of Health Economics*, 2018, 46 (19), 78–90.
- Bonfiglioli, Alessandra**, **Rosario Crinò**, **Harald Fadinger**, and **Gino Gancia**, “Robot Imports and Firm-Level Outcomes,” *Economic Journal (forthcoming)*, 2024.
- Bonnet, Céline** and **Vincent Réquillart**, “Tax incidence with strategic firms in the soft drink market,” *Journal of Public Economics*, 2013, 106, 77–88.
- Borusyak, Kirill** and **Xavier Jaravel**, “Revisiting Event Study Designs, with an Ap-

-
- plication to the Estimation of the Marginal Propensity to Consume," *Working paper*, 2017.
- , – , and **Jann Spiess**, "Revisiting Event Study Designs: Robust and Efficient Estimation," *Working Paper*, 2022.
- Braxton, J. Carter and Bledi Taska**, "Technological Change and the Consequences of Job Loss," *American Economic Review*, 2023, 113, 279–316.
- Britto, Diogo G. C., Paolo Pinotti, and Breno Sampaio**, "The Effect of Job Loss and Unemployment Insurance on Crime in Brazil," *Econometrica*, 2022, 90 (4), 1393–1423.
- Browning, Martin and Eskil Heinesen**, "Effect of job loss due to plant closure on mortality and hospitalization," *Journal of Health Economics*, 2012, 31, 599–616.
- , **Anne Moller Dano, and Eskil Heinesen**, "Job displacement and stress-related health outcomes," *Health Economics*, 2006, 15, 1061–1075.
- Brynjolfsson, Erik and Andrew McAfee**, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: Norton, 2014.
- Callaway, Brantly and Pedro Sant'Anna**, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Case, Anne and Angus Deaton**, "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century," *Proceedings of the National Academy of Sciences*, 2015, 112 (49), 15078–15083.
- Cederlöf, Jonas, Peter Fredriksson, Arash Nekoei, and David Seim**, "Mandatory Advance Notice of Layoff: Evidence and Efficiency Considerations," *Working Paper*, 2021.
- Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer**, "The effect of minimum wages on low-wage jobs," *The Quarterly Journal of Economics*, 2019, 134 (3), 1405–1454.
- Cestone, Giacinta, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica**, "Exploiting Growth Opportunities: The Role of Internal Labour Markets," *Review of Economic Studies*, 2023, 00, 1–41.
- Clainche, Christine Le and Pascale Lengagne**, "The Effects of Mass Layoffs on Mental

-
- Health," *IRDES Working Paper*, 2019, 78.
- Couch, Kenneth and Dana Placzek**, "Earnings Losses of Displaced Workers Revisited," *American Economic Review*, 2010, 100 (1), 572–589.
- DARES**, "Quelles étaient les conditions de travail en 2019, avant la crise sanitaire ?," *DARES Analyses*, 2021, 44.
- **and Pole Emploi**, "Note méthodologique sur l'analyse des tensions," *Document de travail*, 2020.
- de Chaisemartin, Clément and Xavier D'Haultfoeuille**, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, 110 (9), 2964–2996.
- **and –**, "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *Working paper*, 2020.
- Diamond, Peter A**, "A many-person Ramsey tax rule," *Journal of Public Economics*, 1975, 4 (4), 335–342.
- Domini, Giacomo, Marco Grazzi, Daniele Moschella, and Tania Treibich**, "For whom the bell tolls: The firm-level effects of automation on wage and gender inequality," *Research Policy*, 2022, 51 (7), 104533.
- Dubois, Pierre, Rachel Griffith, and Martin O'Connell**, "How well targeted are soda taxes?," *American Economic Review*, 2020, 110 (11), 3661–3704.
- , – , **and –**, "The Use of Scanner Data for Economics Research," *Annual Review of Economics*, 2022, 14, 723–745.
- Eliason, Marcus and Donald Storrie**, "Does Job Loss Shorten Life?," *The Journal of Human Resources*, 2009, 44 (2), 277–302.
- **and –**, "Job loss is bad for your health – Swedish evidence on cause-specific hospitalization following involuntary job loss," *Social Science & Medicine*, 2009, 68, 1396–1406.
- Findeisen, Sebastian, Wolfgang Dauth, and Oliver Schlenker**, "Organized Labor Versus Robots? Evidence From Micro Data," *CEPR Discussion Paper No. 19192.*, 2024.
- Frey, Carl and Michael Osborne**, "The future of employment: How susceptible are jobs to computerisation?," *Technological Forecasting and Social Change*, 2017, 114, 254–

280.

Gerard, François and Joana Naritomi, “Job Displacement Insurance and (the Lack of) Consumption-Smoothing,” *American Economic Review*, 2021, 111 (4), 899–942.

Gihleb, Rania, Osea Giuntella, Luca Stella, and Tianyi Wang, “Industrial Robots, Workers’ Safety, and Health,” *Labour Economics*, 2022, 78, 102205.

Giuntella, Osea, Johannes König, and Luca Stella, “Artificial Intelligence and Workers’ Well-Being,” *IZA Discussion Paper Series*, 2023.

Goldin, Claudia and Lawrence F. Katz, “The Origins of Technology-Skill Complementarity,” *The Quarterly Journal of Economics*, 1998, 113 (3), 693–732.

Goodman-Bacon, Andrew, “Difference-in-differences with variation in treatment timing,” *Journal of Econometrics*, 2021, 225, 254–277.

Goos, Maarten, “The impact of technological progress on labour markets: Policy challenges,” *Oxford Review of Economic Policy*, 7 2018, 34, 362–375.

– **and Alan Manning**, “Lousy and Lovely Jobs: The Rising Polarization of Work in Britain,” *The Review of Economics and Statistics*, 2007, 89, 118–133.

Graetz, Georg and Guy Michaels, “Robots at work,” *The Review of Economics and Statistics*, 12 2018, 100, 753–768.

Griffith, Rachel, Martin O’Connell, and Kate Smith, “Tax design in the alcohol market,” *Journal of Public Economics*, 2019, 172, 20–35.

Gunadi, Christian and Hanbyul Ryu, “Does the rise of robotic technology make people healthier?,” *Health Economics*, 2021, 30, 2047–2062.

Heiss, Florian and Viktor Winschel, “Likelihood approximation by numerical integration on sparse grids,” *Journal of Econometrics*, 2008, 144 (1), 62–80.

Hernnäs, Sofia, “Mortality, Morbidity, and Occupational Decline,” *IFAU Working paper*, 2023.

Humlum, Anders, “Robot Adoption and Labour Market Dynamics,” *Working paper*, 2021.

Hussam, Reshmaan, Erin M. Kelley, Gregory Lane, and Fatima Zahra, “The Psy-

-
- chosocial Value of Employment: Evidence from a Refugee Camp," *American Economic Review*, 2022, 112 (11), 3694–3724.
- Jacobson, Louis, Robert LaLonde, and Daniel Sullivan**, "Earnings Losses of Displaced Workers," *American Economic Review*, 1993, 83 (4), 685–709.
- Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas**, "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration," Tinbergen Institute Discussion Papers 15-110/IV, Tinbergen Institute September 2015.
- Koch, Michael, Ilya Manuylov, and Marcel Smolka**, "Robots and Firms," *Economic Journal*, 2021, 131 (638), 2553–2584.
- Kuhn, Andreas, Rafael Lalive, and Josef Zweimüller**, "The public health costs of job loss," *Journal of Health Economics*, 2009, 28, 1099–1115.
- Lengagne, Pascale**, "Job destruction, disability and mortality," *Working Paper*, 2022.
- Leroutier, Marion and H el ene Ollivier**, "The Cost of Air Pollution for Workers and Firms," *WP*, 2023.
- Lindner, Attila, Bal azs Murak ozy, Bal azs Reizer, and Ragnhild Schreiner**, "Firm-level Technological Change and Skill Demand," *Working Paper*, 2022.
- Lordan, Grace and Eliza-Jane Stringer**, "People versus machines: The impact of being in an automatable job on Australian worker's mental health and life satisfaction," *Economics and Human Biology*, 2022, 46, 101144.
- McFadden, Daniel**, "Econometric models of probabilistic choice," *Structural analysis of discrete data with econometric applications*, 1981, 198272.
- , "Computing Willingness-to-Pay in Random Utility Models," in "Trade, Theory and Econometrics: Essays in Honor of John S. Chipman," London: Routledge, 1999, pp. 253–274.
- **and Kenneth Train**, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 2000, 15 (5), 447–470.
- Nevo, Aviv**, "Measuring market power in the ready-to-eat cereal industry," *Economet-*

-
- rica*, 2001, 69 (2), 307–342.
- Obesity Evidence Hub**, “Countries that have taxes on sugar-sweetened beverages (SSBs),” Technical Report 2022. Available online at <https://www.obesityevidencehub.org.au/collections/prevention/countries-that-have-implemented-taxes-on-sugar-sweetened-beverages-ssbs>, accessed March, 2024.
- O’Connell, Martin and Kate Smith**, “Optimal Sin Taxation and Market Power,” *American Economic Journal: Applied Economics*, 2024, 16 (4), 34—70.
- OECD**, “Health at a Glance 2019: OECD Indicators,” *Éditions OCDE*, 2019.
- Petrin, Amil and Kenneth Train**, “A Control Function Approach to Endogeneity in Consumer Choice Models,” *Journal of Marketing Research*, 2010, 47 (1), 3–13.
- Pigou, Arthur**, *The economics of welfare*, 4th ed., Macmillan, 1934.
- Public Health England**, “Results of the National Diet and Nutrition Survey (NDNS) rolling programme for 2014 to 2015 and 2015 to 2016,” Technical Report 2018. Available online at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/699241/NDNS_results_years_7_and_8.pdf, accessed March, 2019.
- Rabaté, Simon and Julie Rochut**, “Employment and Substitution Effects of Raising the Statutory Eligibility Age in France,” *PSE Working Paper*, 2017, 46.
- Radloff, Lenore Sawyer**, “The CES-D scale: A self report depression scale for research in the general population,” *Applied Psychological Measurements*, 1977, 1 (3), 385–401.
- Rege, Mari, Kjetil Telle, and Mark Votruba**, “The effect of plant downsizing on disability pension utilization,” *Journal of the European Economic Association*, 2009, 7 (4), 754–785.
- Restrepo, Pascual**, “Automation: Theory, Evidence, and Outlook,” *NBER working paper*, 2023, 31910.
- Revelt, David and Kenneth Train**, “Mixed Logit with Repeated Choices: Households’ Choices of Appliance Efficiency Level,” *Review of Economics and Statistics*, 1998, 80 (4), 647–657.

- Rischbieth, Amanda, George Serafeim, and Katie Trinh**, "Accounting for product impact in the consumer-packaged foods industry," *Harvard Business School Working Paper*, 2020, 21-051.
- Roulet, Alexandra**, "The Causal Effect of Job Loss on Health: The Danish Miracle?," *Working Paper*, 2017.
- Schaller, Jessamyn and Ann Huff Stevens**, "Short-run effects of job loss on health conditions, health insurance, and health care utilization," *Journal of Health Economics*, 2015, 43, 190–203.
- Schmieder, Johannes F., Till von Wachter, and Jörg Heining**, "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany," *American Economic Review (forthcoming)*, 2022, 90 (4), 1393–1423.
- Schmitz, Hendrik**, "Why are the unemployed in worse health? The causal effect of unemployment on health," *Labour Economics*, 2011, 18, 71–78.
- Small, Kenneth A and Harvey S Rosen**, "Applied welfare economics with discrete choice models," *Econometrica*, 1981, pp. 105–130.
- Sullivan, Daniel and Till von Wachter**, "Job Displacement and Mortality: An Analysis using administrative Data," *Quarterly Journal of Economics*, 2009, 124 (2), 1265–1306.
- Sun, Liyang and Sarah Abraham**, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, 225 (2), 175–199.
- Webb, Michael**, "The Impact of Artificial Intelligence on the Labor Market," *Working Paper*, 2020.
- World Bank**, "Digital Dividends," *World Development Report 2016*, 2016.
- Zins, Marie, Marcel Goldberg, and Constances-team**, "The French CONSTANCES population-based cohort: design, inclusion and follow-up," *European Journal of Epidemiology*, 2015, 30 (12), 1317–1328.



List of Figures

- 1 Timeline of industrial revolutions 30
- 2 Trends in robot density 31
- 3 Evolution of the prevalence of obesity and overweight 34
- 4 Frise chronologique des révolutions industrielles 53
- 5 Évolution de la densité de robots 54
- 6 Evolution de la prévalence de l’obésité et du surpoids 58

- 1.1 Employment status and time to event 93
- 1.2 Effect of job loss on economic outcomes 98
- 1.3 Effect of job loss on mental health 100
- 1.4 Heterogeneity (mental health-related drugs) 101
- 1.5 Individual risk on the labor market and mental health 104
- 1.A.1 Map of targeted zones by Constances 108
- 1.A.2 Participation to survey 108
- 1.D.1 Timing of employees’ exit before closure 111
- 1.D.2 Composition of the treated group 111
- 1.F.1 Trends in earnings 113
- 1.F.2 Robustness checks using standard TWFE and only ever-treated individuals 115
- 1.G.1 Heterogeneity (economic outcomes) 116
- 1.H.1 Effect of job loss on physical health 117
- 1.H.2 Heterogeneity (physical health) 120
- 1.H.3 Effect of job loss on disability benefits and sick leaves 121
- 1.H.4 Heterogeneity (health insurance) 123

1.I.1	Predicted reemployment probabilities and health outcomes	125
2.1	Evolution of the stock of tangible fixed assets and automation capital by proxy	139
2.1	Macro trends in automation imports	140
2.2	Lumpiness of imports of automation technologies	142
2.3	Evolution of the workforce and the stock of automation technologies . .	143
2.1	Event studies at the firm level - results on automation	153
2.2	Event studies at the firm level - main results	154
2.1	Event studies at the worker level - labor market outcomes	160
2.2	Event studies at the worker level - mental health outcomes	161
2.3	Event studies at the worker level - physical health outcomes	162
2.4	Heterogeneity	163
2.5	Event studies at the worker level - health insurance	164
2.A.1	Mechanisms	172
2.B.1	Map of targeted zones by Constances	175
2.C.1	Macro trends in automation imports	180
2.C.2	Main origin countries of automation technologies and industrial robots .	180
2.C.3	Imports of automation technologies by French industry	181
2.C.4	Stock of automation capital by sector	181
2.C.5	Relative size of the import according to its rank - industrial robots	182
2.C.6	Relative size of the import according to its rank - by sector	182
2.C.7	Workers' exposure to automation	183
2.D.1	Distribution of the log changes	185
2.D.2	Evolution of the p90 of log changes in automation stock over time	186
2.D.3	Employment status and time to event	188
2.E.1	Event studies at the firm level - results on automation	189
2.E.2	Event studies at the firm level (results on employment) - comparison of different proxies for the stock of goods	190
2.E.3	Event studies at the firm level (results on employment) - comparison of different measures	191

2.E.4 Event studies at the firm level (results on sales) - comparison of different measures	193
2.E.5 Event studies at the firm level (results on the share of blue-collar workers) - comparison of different measures	195
2.E.6 Event studies at the firm level - treatment: first import of industrial robots	197
2.E.7 Event studies at the firm level (results on employment) - comparison of different specifications	199
2.E.8 Event studies at the firm level (results on sales) - comparison of different specifications	200
2.E.9 Event studies at the firm level (results on the share of blue-collar workers) - comparison of different specifications	201
2.F.1 Heterogeneity (labor market outcomes)	202
2.F.2 Event studies at the worker level - mental health outcomes	203
2.F.3 Event studies at the worker level - physical health outcomes	204
2.F.4 Heterogeneity (mental health outcomes)	205
2.F.5 Heterogeneity (general health outcomes)	207
2.F.6 Heterogeneity (health insurance)	209
3.1 Distribution of the sugar content of the non-alcoholic beverage supply .	217
3.1 Comparison of tiered tax designs	229
3.2 Effects on prices and market shares	231
3.3 Welfare change in linear taxation with tax rate	233
3.A.1 Distribution of the sugar content of the non-alcoholic beverage supply .	240
3.A.2 Purchase of non-alcoholic beverages across Kantar households	242
3.A.3 Sugar purchases across Kantar households	242
3.A.4 Design of the SDIL	243
3.D.1 Comparison of tiered tax designs (without firms' strategic behavior) . .	248
3.D.2 Welfare variation according to the location of thresholds	249
3.D.3 Welfare variation according to the location of threshold ($k = 1$)	249
3.D.4 Welfare variation according to the location of thresholds ($k = 2$)	250



List of Tables

- 1 Sugar taxes on SSBs across Europe 36
- 2 Taxes sur les boissons sucrées en Europe 60
- 1.1 Summary statistics (year $c-2$) 92
- 1.B.1 Requirements to validate quarters in each status 109
- 1.C.1 Definition of health outcomes 110
- 1.D.1 Explanation of the sample size of the treated group 110
- 1.E.1 Summary statistics on industry and socio-professional category (year -2) 112
- 1.F.1 Robustness checks on the specification 113
- 1.H.1 Disaggregated outcomes for doctors’ visits 118
- 1.H.2 Disaggregated outcomes for hospital admissions 119
- 1.H.3 Disaggregated outcomes for long-term conditions 122
- 2.1 Sample size - Firms & Workers 147
- 2.2 Summary statistics (year 2007) 149
- 2.3 Summary statistics (year $c - 2$) 151
- 2.A.1 Literature review on technological change & labor market outcomes . . 173
- 2.A.2 Literature review on technological change & health outcomes 174
- 2.B.1 Requirements to validate quarters in each status 175
- 2.B.2 Definition of health outcomes 176
- 2.B.3 HS-2012 product codes referring to automation technologies and robots
[\[Acemoglu and Restrepo, 2022\]](#) 178
- 2.D.1 Sample composition 187
- 3.1 Welfare variation decomposition 230

3.2	Welfare variation decomposition without firms' strategic behavior	232
3.A.1	Household characteristics in Kantar panel	238
3.A.2	Summary of the market definition	238
3.A.3	Market structure	239
3.A.4	Average market share, price and sugar content of each sub-category . . .	241
3.A.5	Household purchases	243
3.B.1	Results on price equation	244
3.B.2	Estimates of the random coefficient logit model	246
3.B.3	Own-price elasticities by household characteristics	246
3.B.4	Own-price elasticities by sugar content and household characteristics . .	247
3.C.1	Elasticities and margins by firms	247
3.C.2	Margins by sugar content	247

Titre : Essais en Economie de la Santé et du Travail

Mots clés : Consommation de soins, Santé mentale et physique, Perte d'emploi, Automatisation, Taxes nutritionnelles, Design optimal

Résumé : Cette thèse explore deux déterminants clés de la santé : les conséquences d'événements professionnels et la conception de politiques fiscales nutritionnelles. Le premier objectif est d'améliorer la compréhension des liens entre certains chocs liés à l'emploi et la santé des individus. Le premier chapitre examine les effets de la perte involontaire d'emploi sur la consommation de soins. Les pertes d'emplois exogènes sont identifiées par le biais de fermetures d'établissements, et des estimations causales sont obtenues en comparant un groupe de travailleurs exposés à un groupe similaire de travailleurs non concernés. Les résultats montrent que la perte d'emploi a des impacts négatifs à long terme sur la santé mentale, tandis que les effets sur la santé physique sont moins persistants. Le deuxième chapitre étudie les effets des transformations technologiques sur le lieu de travail, en particulier l'automatisation et la robotique, sur la santé des travailleurs. L'automatisation au niveau de l'entreprise est mesurée à l'aide de proxys tels que le stock d'équipement industriel et de machines, ainsi que les importations de biens d'automatisation et de robots industriels. Les analyses au niveau des entre-

prises révèlent que l'effet de productivité compense l'effet de déplacement : l'automatisation est associée à une augmentation du nombre d'employés, en particulier parmi les ouvriers. Les travailleurs en place augmentent leur consommation d'antidépresseurs et d'anxiolytiques, bien qu'aucun changement significatif concernant les soins liés à la santé physique ne soit observé. L'autre objectif scientifique est d'approfondir la compréhension des politiques fiscales conçues pour lutter contre l'obésité. Le troisième chapitre évalue les effets des taxes nutritionnelles sur les boissons non alcoolisées sur le bien-être global, en tenant compte de l'utilité des consommateurs, des entreprises, du gouvernement et des externalités générées par une consommation excessive de sucre. L'étude propose des conceptions optimales de taxes progressives, dérivées empiriquement, qui sont ensuite comparées aux conceptions optimales suggérées par la littérature théorique. Cette thèse, basée sur une approche empirique, s'appuie sur un ensemble de données riche et diversifié : des données administratives (au niveau entreprises et individus), des enquêtes et des données d'achat des ménages.

Title : Essays in Health and Labor Economics

Keywords : Healthcare use, Mental and physical health, Job loss, Automation, Sugar taxes, Optimal design

Abstract : This thesis explores two key determinants of health : the consequences of career events and the design of nutritional fiscal policies. The first objective is to enhance the understanding of the links between certain employment shocks and individuals' health. The first chapter examines the effects of involuntary job loss on healthcare consumption. Exogenous job losses are identified through plant closures, and causal estimates are obtained by comparing a group of displaced workers to a similar group of non-displaced workers. The findings show that job loss has long-term negative impacts on mental health, while the effects on physical health are less persistent. The second chapter investigates the effects of technological transformations in the workplace, particularly automation and robotics, on workers' health. Automation at the firm level is measured using proxies such as the stock of industrial equipment and machines, as well as imports of automation goods and industrial robots. Firm-level analyses reveal that the productivity effect

offsets the displacement effect : automation is associated with an increase in the number of employees, especially among blue-collar workers. Incumbent workers increase their consumption of antidepressants and anxiolytics, though no significant changes in healthcare related to physical health are observed. The other scientific aim is to deepen the understanding of fiscal policies designed to combat obesity. The third chapter evaluates the effects of nutritional taxes on non-alcoholic beverages on overall welfare, considering the utility of consumers, firms, the government, and the externalities generated by excessive sugar consumption. The study provides empirically derived optimal designs for tiered taxes, which are then compared to optimal designs suggested by theoretical literature. This thesis, based on an empirical approach, draws on a rich and diverse set of data : administrative data (both at the firm and individual levels), surveys, and household purchase data.