



**HAL**  
open science

# Trois essais empiriques en économie de la santé et du travail

Alexandre Cazenave-Lacroutz

► **To cite this version:**

Alexandre Cazenave-Lacroutz. Trois essais empiriques en économie de la santé et du travail. Economies et finances. Université Paris sciences et lettres, 2021. Français. NNT : 2021UPSLD007 . tel-03546716

**HAL Id: tel-03546716**

**<https://theses.hal.science/tel-03546716>**

Submitted on 28 Jan 2022

**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.



**THÈSE DE DOCTORAT**  
**DE L'UNIVERSITÉ PSL**

Préparée à l'Université Paris-Dauphine

**Trois essais empiriques**  
**en économie de la santé et du travail**

Soutenue par

Alexandre

**Cazenave-Lacroutz**

Le 2 juin 2021

École doctorale n°ED543

**Ecole doctorale SDOSE**

Spécialité

**ECONOMIE**

Composition du jury :

M. Andrew CLARK

Directeur de recherche au CNRS,  
PSE - École d'économie de Paris

*Rapporteur*

M. Joachim WINTER

Professeur des universités,  
Université Louis-et-Maximilien de Munich

*Rapporteur*

Mme. Brigitte DORMONT

Professeure des universités,  
Université Paris Dauphine-PSL

*Présidente du jury*

M. Didier Blanchet

Administrateur de l'Insee

*Examineur*

Mme. Eve CAROLI

Professeure des universités,  
Université Paris Dauphine-PSL

*Directrice de thèse*



# Remerciements

Au-delà des remerciements propres à chaque chapitre, des remerciements plus généraux s'imposent.

Je tiens en premier lieu à témoigner ma profonde gratitude envers **Eve Caroli** pour avoir accepté de diriger cette thèse, mais aussi, et surtout, pour sa direction. Elle fut non seulement directe et exigeante quand ce que je lui présentais témoignait trop visiblement de mon inexpérience ou d'un manque de finition, mais aussi inspirante et présente, notamment dans les moments de doute. Elle m'a laissé l'autonomie nécessaire pour le bon développement de cette thèse tout en me donnant les clés pour avancer droit au but et en m'enseignant par l'exemple comment raisonner en économiste. J'émetts le souhait de continuer à apprendre à son contact bien au-delà de cette thèse !

Je remercie également **Brigitte Dormont** et **Didier Blanchet**, qui ont accepté de prendre part au jury de cette thèse malgré leur participation préalable à pas moins de trois comités de suivi de thèse. Ils sont aussi responsables d'améliorations notables dans les chapitres, et furent des critiques patients et encourageants.

Des remerciements appuyés se doivent aussi d'être formulés à l'attention de **Joachim Winter** et d'**Andrew Clark**, pour avoir accepté d'être rapporteurs de cette thèse. Leurs seules critiques à la pré-soutenance ont permis d'une part des clarifications importantes dont bénéficieront tous les lecteurs de cette thèse, et d'autre part certaines avancées pour des articles que je pensais plus aboutis.

---

Je n'oublie pas de citer mes deux co-auteurs, respectivement **Alexandre Godzinski** pour le premier chapitre, et **Engin Yilmaz**, pour le second : ce fut un plaisir et un enrichissement de travailler avec eux sur ces projets, et j'ai pour ma part beaucoup bénéficié de leur expérience.

Je témoigne ensuite, avec amitié, de toute ma gratitude envers **Fanny Godet** et **Tomáš Jagelka** pour nos échanges réguliers qui m'ont aidé dans la recherche d'un ton juste. Tomáš a également lu et relu tous les articles, parfois à (très) brève échéance, critiquant avec justesse et précision, tout en prodiguant des conseils pertinents. Se sont aussi fortement dévoués aux exercices finaux de relecture mon amie **Alice Halys** et ma femme **Hélène**, ainsi que mon frère **Laurent Cazenave-Lacrouz** et mon ami **Antoine Halys**. Je les en remercie aussi.

Cette thèse n'aurait pas non plus été possible sans les institutions où elle s'est préparée. Je souhaite citer avant tout l'institution de crainte d'oublier des personnes qui m'y ont aidé : **l'Université Paris Dauphine et le LEDa** qui ont permis le déroulement harmonieux de cette thèse dans des conditions quelque peu atypiques – merci à **Henri Cazenave-Lacrouz** pour y avoir procédé par deux fois à ma réinscription ; **l'Insee** ensuite, et plus particulièrement le **Département des Etudes Economiques**. Je souhaiterais notamment y remercier nominativement **Corinne Prost**, qui m'a convaincu d'y venir, **Malik Koubi** qui m'y a accueilli et encouragé, **Sébastien Roux**, **Didier Blanchet** (en sa qualité de Directeur de la DESE) et **Lionel Wilner**, pour leurs relectures attentives et leurs suggestions constructives. Je remercie également **la Drees**, et notamment le **Bureau des Etablissements de Santé** (le BES), où **Gwennaëlle Brillhault** puis **Fabien Toutlemonde**, et **toute l'équipe** m'ont complètement intégré en stage de fin d'étude, puis en collaboration, dans une ambiance de travail tout aussi studieuse que joyeuse. Je remercie aussi à la **LMU** le professeur **Joachim Winter**, pour m'avoir accueilli en *visiting student* pendant un an, ainsi que **Fabien**

---

**Stürmer-Heiber** et **Felix Hagemeister** (et les autres doctorants de la chaire) pour leur accueil chaleureux pendant cette année purement académique. Enfin, je remercie le **Crest**, non seulement pour le support matériel indispensable qu’il m’a fourni au travers de mon affiliation ces dernières années, mais aussi pour les échanges précieux que j’ai eus avec certains de ses membres ; outre Tomáš, **Pierre Cahuc** en particulier a inspiré le premier chapitre de la thèse, et **Emilie Sartre** a été l’une des premières lectrices du troisième chapitre.

Enfin, je me suis marié pendant cette thèse, et j’ai eu le bonheur d’avoir avec ma femme deux enfants, certes actifs, mais dont la seule vue me fait aussi sourire et qui tentent parfois de reposer leur « Papa-fatigué » par un câlin. Merci aussi et surtout à mon extraordinaire **Hélène** de m’avoir continuellement soutenu dans ce long trajet quand bien même il nous aura coûté, sur une période prolongée, tous nos temps libres et beaucoup d’efforts.

En fils bien-aimant – et en jeune papa – je saisis finalement cette occasion pour remercier plus généralement mes parents pour tout ce qu’ils ont fait, font encore et transmettent à leurs enfants, avec pour cette thèse, une mention particulière à ma mère qui m’aura demandé tous les mois pendant quatre ans : « Comment avance ta thèse ? ».

La voici enfin, cher Lecteur ! – merci de vous être décidé à la lire.



# Sommaire

<b>Remerciements</b>	<b>1</b>
<b>Sommaire</b>	<b>5</b>
<b>Introduction Générale</b>	<b>9</b>
Les questions . . . . .	10
Les méthodes . . . . .	12
Recommandations de politiques publiques . . . . .	17
Bibliographie . . . . .	27
<b>1 Les effets d'un jour de carence pour arrêt maladie sur les absences pour raison de santé dans la fonction publique de l'État française</b>	<b>31</b>
<i>Lower sick leave cover, fewer health-related work absences ?</i>	
1.1 <i>Introduction</i> . . . . .	34
1.2 <i>Institutional framework, data, and descriptive statistics</i> . . . . .	37
1.3 <i>Empirical strategy</i> . . . . .	52
1.4 <i>Results</i> . . . . .	54
1.5 <i>Robustness tests</i> . . . . .	61
1.6 <i>Conclusion</i> . . . . .	67
1.7 <i>References</i> . . . . .	69
1.8 <i>Appendices</i> . . . . .	73



**2 Incitations par les prix, monitoring et pratiques médicales des hôpitaux 79**

*Can price incentives or the threat of monitoring induce hospitals to change their medical practices ? Evidence from two French policies.*

2.1	<i>Introduction</i>	82
2.2	<i>The single tariff policy and the prior approval policy</i>	87
2.3	<i>Data and the studied treatments</i>	92
2.4	<i>The Empirical strategy</i>	96
2.5	<i>Results</i>	106
2.6	<i>Conclusion</i>	128
2.7	<i>References</i>	132
2.8	<i>Glossary</i>	137
2.9	<i>Appendices</i>	143

**3 Temps de transport domicile-travail, absentéisme et présentéisme en France 173**

*Does commuting time affect absenteeism and presenteeism in France ?*

3.1	<i>Introduction</i>	176
3.2	<i>Data and variables</i>	180
3.3	<i>Empirical strategy</i>	184
3.4	<i>Results</i>	187
3.5	<i>The intermediation of health in the studied relationships</i>	219
3.6	<i>Conclusion</i>	227
3.7	<i>References</i>	230
3.8	<i>Appendices</i>	235

<b>4 Conclusion générale</b>	<b>259</b>
4.1 Quels bénéfices/coûts pour les politiques publiques étudiées? . . . .	259
4.2 Quelles nouvelles question de recherches? . . . . .	265
4.3 Bibliographie . . . . .	269
<b>Table des matières</b>	<b>271</b>



# Introduction Générale

## Quelles mesures pour les politiques publiques ? <sup>1</sup>

A l’instar des physiciens dont les recherches théoriques et expérimentales sont à la base des artefacts des ingénieurs, l’une des visées des travaux des économistes peut être la meilleure compréhension, et *in fine* l’amélioration, des politiques publiques par ceux qui les décident et les mettent en oeuvre. Les trois chapitres de cette thèse y contribuent, chacun à son échelle et sur un sujet distinct.

Les trois chapitres se rapportent cependant tous à la santé et au travail d’une part, et aux incitations d’autre part. La santé et le travail d’une part, parce que j’étudie soit des facteurs pouvant influencer les arrêts maladies ou le présentéisme des travailleurs (dans les chapitres 1 et 3), soit comment la rémunération de ses acteurs peut altérer la production de soins (au chapitre 2). Les incitations d’autre part, soit parce que les deux politiques publiques étudiées visaient à modifier les incitations économiques des agents (au chapitre 1 en modifiant les règles d’indemnisation en cas d’arrêt maladie, au chapitre 2 en modifiant les tarifs rémunérant les hôpitaux pour les séjours hospitaliers), soit parce que les incitations économiques sont l’un des fondements théoriques de la relation étudiée (au chapitre 3).

Cette introduction s’articulera en trois temps : Je m’attache d’abord à montrer en quoi l’objet, la question centrale, de chaque chapitre, vise à la mesure de l’effet des politiques publiques, et à l’amélioration des mesures de politiques

---

1. Je remercie pour leurs commentaires et suggestions Eve Caroli, Alice Halys, Hélène Hartmann, et Tomáš Jagelka.

publiques. J'explicite ensuite pourquoi les méthodes employées diffèrent autant. Enfin, je synthétise les principaux résultats et cherche à en tirer des recommandations raisonnées pour les politiques publiques.

\* \* \*

## Les questions

Le premier chapitre, co-écrit avec Alexandre Godzinski, s'intitule en anglais «*Lower sick leave cover, fewer health-related work absences ?*», ce qui se traduirait par «Est ce qu'une plus faible couverture des arrêts maladies entraîne moins d'absences pour raison de santé?». Nous y identifions les effets attribuables à l'introduction d'un jour de carence pour les arrêts maladies, mis en place en France en 2012-2013 dans la fonction publique.

Le deuxième chapitre, co-écrit avec Engin Yilmaz, s'intitule en anglais «*Can price incentives or the threat of monitoring induce hospitals to change their medical practices ? Evidence from two French policies.*», à savoir : «Est ce que les incitations tarifaires ou la menace de contrôles peuvent faire changer les pratiques médicales des hôpitaux? Des résultats basés sur deux politiques françaises.» Nous y évaluons si certaines incitations tarifaires visant les séjours en chirurgie ambulatoire (c'est à dire quand le patient peut rentrer chez lui le jour même de son opération chirurgicale) ou une procédure de mise sous accord préalable par l'Assurance Maladie (des séjours non réalisés en chirurgie ambulatoire) conduit en effet à une augmentation de la pratique de la chirurgie ambulatoire, comme escompté.

Le troisième chapitre s'intitule en anglais «*Does commuting time affect absenteeism and presenteeism in France ?*». La question posée est en effet : «Est-ce que le temps de transport domicile-travail influe sur l'absentéisme et le présentéisme en France?»; où l'absentéisme désigne le nombre de jours d'arrêts

---

maladies sur une année, et le présentéisme le nombre de jour sur une année où le salarié a travaillé alors qu'il était malade.

Les chapitres 1 et 2 sont donc des **évaluations empiriques, *ex-post*, de politiques publiques**. Plus précisément, elles visent chacune à mesurer si une modification précise d'une politique publique a bien conduit aux effets visés par ce changement. Si non, il s'agit d'identifier, ou à défaut, de proposer des pistes d'explications quant aux causes du décalage par rapport aux attentes initiales, et éventuellement d'étudier si d'autres indicateurs importants sont également affectés. De telles études peuvent servir à éclairer l'action publique, par exemple aux fins de corriger ou d'améliorer des politiques existantes, et leur intérêt semble grandissant en France (voir par exemple Bozio (2014), Givord (2014), CAE (2013)). Plus généralement, de telles évaluations peuvent évidemment être utilisées en dehors du cadre spécifique où elles ont été produites, en témoigne à titre d'exemple proéminent les travaux de la Commission d'étude des effets de la loi pour la croissance et l'activité<sup>2</sup> dont les 5 études mobilisaient abondamment les évaluations de politiques publiques à l'étranger pour éclairer les impacts que des changements similaires pourraient avoir en France (voir Bozio et al. (2015)).

En comparaison, le chapitre 3 peut **paraître au premier abord plus éloigné de l'action publique**, puisqu'il s'agit ici d'identifier l'impact du temps de trajet domicile-travail sur l'absentéisme et le présentéisme. Mais les politiques publiques en terme de transport, de logement, de travail peuvent affecter le temps de trajet domicile-travail : par la construction d'infrastructures ou de réseaux ; par les autorisations de constructions ; par les exonérations d'impôts sur les frais de transport (entre autres). Selon qu'une relation est avérée entre de telles politiques et des variables telles que l'absentéisme et le présentéisme, le chapitre 3 permet d'en tenir compte quantitativement dans les analyses coûts-bénéfices ; ou

---

2. Cette commission, composée de 10 experts indépendants (majoritairement des professeurs d'économie), visait à fournir aux parlementaires et aux citoyens une évaluation *ex-ante* des effets attendus de cette loi (votée en 2015).

dans les modèles structurels ou modèles de microsimulation **permettant d'évaluer *ex-ante* les effets attendus des politiques**. Ce qui permet d'ajuster les paramètres des mesures envisagées.

Plus généralement, les trois chapitres visent aussi à améliorer les politiques publiques à plus long terme, en ce qu'ils contribuent aussi chacun (à leur échelle) à une meilleure compréhension de mécanismes théoriques. S'il ne le fait pas immédiatement par lui-même, un chercheur empirique montrant que les prédictions des mécanismes théoriques sont parfois invalides peut espérer que cela générera ultérieurement des raffinements de la théorie économique, et *in fine* de meilleures propositions de politiques publiques. Pour ne donner ici qu'un exemple, la littérature économique empirique préalable au premier chapitre montre assez explicitement qu'une diminution de l'indemnisation des arrêts maladies peut conduire à une réduction de l'absentéisme ; sans en présenter tous les résultats, le premier chapitre (comme d'autres articles sur des dispositifs semblables au jour de carence) met cependant en évidence qu'une diminution de l'indemnisation des arrêts maladies n'a pas toujours les effets attendus ; la façon dont cette baisse s'opère pourrait aussi compter.<sup>3</sup> Si cette dernière hypothèse se confirmait, des modèles théoriques plus complexes que le modèle classique d'arbitrage travail-loisir pourraient affiner notre compréhension des formes "optimales" de compensation des arrêts maladies.

\* \* \*

## Les méthodes

Pour apporter ces contributions aux mesures des effets des politiques publiques, des méthodes différentes, adaptées aux situations rencontrées par chaque article, ont été utilisées.

---

3. C'est à dire : les résultats semblent assez différer entre une baisse uniforme du taux de remplacement avec la durée, et une baisse concentrée en début d'arrêt maladie.

---

**Les chapitres visent tous trois à identifier une inférence causale**, c'est à dire à estimer l'effet propre à une variable focale d'intérêt (ici une réforme, là le temps de trajet domicile travail) sur une variable dépendante d'intérêt (ici l'absentéisme ou le présentéisme du fait de la santé, là le taux de chirurgie ambulatoire). Par effet propre, il faut entendre : non entaché de biais divers liés par exemple à des corrélations avec des variables non présentes dans les données,<sup>4</sup> notamment lié à des effets de sélection (si les bénéficiaires d'une mesure ne sont pas directement comparables à ceux qui n'en ont pas bénéficié), etc.

Comme le théorise Rubin (1974), **on peut voir dans une expérience aléatoire contrôlée un cadre idéal pour une analyse empirique**. Similairement à un essai médical où un groupe de patients est tiré au hasard parmi un groupe de patients potentiels pour recevoir un traitement, l'effet moyen du traitement s'estime simplement par la différence entre les patients traités et les patients non-traités. Si cette approche se heurte en réalité à des difficultés pratiques et conceptuelles en recherche médicale, et plus encore en sciences sociales (Heckman, 2020), sa référence permet surtout de mettre en avant la comparaison d'un groupe "testé" avec un groupe "témoin", comparable parce que la procédure de sélection n'a pas reposé sur des variables inobservées. Ainsi, **aucun article de la thèse n'étudie une expérience aléatoire contrôlée**. On n'ose pas imaginer les moyens réglementaires et financiers nécessaires à changer arbitrairement l'indemnisation des arrêts maladies ou même les distances domicile-travail d'un groupe de salariés cobayes! Mais, du fait de méthodes popularisées en économie lors des trois dernières décennies, ce n'est aussi peut-être pas nécessaire. Sont ainsi respectivement utilisées des méthodes de différence de différences, de contrôle synthétique, et de régression à effet fixe (avec prise en compte, au maximum, de ce qui peut générer

---

4. De façon caricaturale, si les personnes les plus motivées par leur travail tendent à bénéficier des régimes d'indemnisation maladie les plus généreux, et sont aussi les moins absentes pour maladie, alors, si l'économètre n'observe pas la motivation des salariés dans ces données, une simple analyse *toutes choses égales par ailleurs* pourrait l'amener à conclure qu'une indemnisation plus généreuse a conduit à moins d'absences pour maladie, et ce même si l'indemnisation n'avait pas vraiment d'effet sur les absences pour maladie. En cas d'effet propre positif d'un indemnisation plus généreuse sur les absences pour maladie, ce biais négatif peut aussi conduire à ne pas observer d'effet significatif d'une indemnisation plus généreuse.



de l'endogénéité).

**Le premier chapitre repose ainsi sur une méthode de différence de différences** : il s'agit de comparer la variable d'intérêt entre un groupe "traité" (ici les fonctionnaires de la fonction publique de l'État) par rapport à un groupe "contrôle" (ici les employés en Contrat à Durée Indéterminée (CDI) du secteur privé), après et avant l'introduction de la mesure qui n'affecte que le groupe traité (ici la mise en place de 2012 à 2014 d'une journée de carence pour l'indemnisation des arrêts maladies). On applique une telle méthode de préférence à une simple comparaison sur le groupe traité entre après et avant l'introduction de la mesure pour éviter tout effet temporel qui introduirait un biais dans une méthode par simple différence. Sous l'hypothèse de tendance commune, à savoir qu'en l'absence de la mesure sur le groupe traité et compte tenu des variables observées, l'évolution avant/après aurait été la même entre le groupe traité et le groupe contrôle, la méthode de différence de différences permet d'éviter ce biais.

La politique étudiée dans ce premier chapitre n'est donc pas une expérience aléatoire contrôlée, mais elle possède tout de même des similarités telles qu'on peut évoquer une "expérience naturelle" au prix d'une restriction sur les individus considérés. Lorsqu'on se concentre en effet sur la fonction publique de l'État, la politique ne s'applique pas, en premier lieu, suivant des caractéristiques individuelles, mais est appliquée abruptement à tous les individus d'un secteur ; pour lesquels nous prétendons, en second lieu, pouvoir trouver des individus "contrôles". Le premier point nous a contraint à exclure la fonction publique territoriale de notre analyse, de nombreuses anecdotes témoignant en effet d'une application très hétérogène de la mesure (et inobservée) dans la fonction publique territoriale ; l'exclusion de la fonction publique hospitalière, qui avait des tendances annuelles très distinctes des autres secteurs avant 2012, est également nécessaire pour rendre crédibles les CDI du secteur privé comme "groupe de contrôle". Cette crédibilité est aussi étayée par les nombreuses statistiques descriptives et un test placebo dans la période précédant la mise en place de la mesure. Enfin, au plan méthodologique, il est

---

également très intéressant de pouvoir observer à la fois l'introduction de la mesure en 2012 et sa suppression en 2014. Même si des effets symétriques ne sont pas nécessaires, obtenir des effets de signe opposé contribue à la crédibilité des effets estimés.

**Le deuxième chapitre utilise une méthode de construction de "contrôles synthétiques"**, reprenant une méthode mise en avant par Abadie, Diamond et Hainmueller (2010). La politique étudiée est loin d'être une expérimentation aléatoire contrôlée : d'une part, le traitement que l'on voulait initialement étudier, une incitation tarifaire, a ainsi bénéficié à plusieurs années d'intervalles à différents groupes de pathologies (plus précisément : de "Groupe Homogène de Malade" ou GHM), **sélectionnés suivant des critères différents et pour l'essentiel inobservés** ; d'autre part, ces groupes diffèrent fortement les uns des autres en niveau et en tendance de la variable d'intérêt (le taux de chirurgie ambulatoire). On s'inspire malgré tout de cet idéal de l'expérimentation aléatoire contrôlée pour essayer d'obtenir un effet causal du traitement. Bien que l'on observe quel GHM a bénéficié de l'incitation tarifaire, il n'était pas possible de dégager aisément un groupe de "contrôle", pour lequel une hypothèse de tendances communes en l'absence de la politique serait crédible : pour essayer malgré tout d'obtenir une inférence causale, l'objet de la méthode du contrôle synthétique est alors de chercher à construire un contrôle le plus crédible possible, en ce qu'il est créé pour reproduire le niveau et l'évolution passée des variables pertinentes.<sup>5</sup> Il s'agit des taux de chirurgie ambulatoire avant la mise en place des incitations, mais aussi d'un taux de potentiel de chirurgie ambulatoire établi par des sociétés savantes médicales. Il semble assez plausible qu'un contrôle synthétique partageant ces observables ait une évolution qui aurait été celle du GHM incité en l'absence de l'incitation. Des tests d'inférence exact de Fisher et des tests de robustesses permettent également de limiter les "faux positifs".

---

5. En pratique, on définit les variables d'intérêt de ce contrôle synthétique comme une pondération (sous contrainte) des GHMs non-incités. Et le jeu de poids retenu pour construire le contrôle synthétique est le jeu de poids qui permet de minimiser l'écart (avant l'incitation) aux variables pertinentes choisies.

Contrairement à l'article d'Abadie, Diamond et Hainmueller (2010), la méthode apparaît d'autant plus complexe que la mesure étudiée s'est mise en place en plusieurs vagues d'incitation, et en parallèle d'une autre politique ayant la même visée (la Mise sous Accord Préalable ou MSAP, également étudiée), qui plus est définie suivant une autre classification de séjours. La méthode est cependant aussi utile du fait de la complexité du dispositif à évaluer : **de par son caractère naturellement "algorithmique", la construction des contrôles synthétique donne l'opportunité d'étudier les GHMs incités de la façon la plus exhaustive possible**, au lieu de sélectionner préalablement quelques cas à étudier et auxquels seraient appliqués des techniques *ad hoc*.

Économétriquement, **le troisième chapitre utilise de simples régressions à effet individuel** pour identifier les éventuels impacts du temps de trajet domicile-travail sur les arrêts maladies et le présentéisme, mais aussi sur la santé auto-déclarée et le bien-être psychologique. Il ne s'agit à l'évidence pas d'une expérience aléatoire contrôlée, mais il peut cependant aussi être rapproché des quasi-expériences. **Une spécificité des études portant sur le trajet domicile-travail est en effet qu'il est possible, avec certaines restrictions, d'envisager des cas s'apparentant à une expérience naturelle.** Par exemple, si une entreprise déménage ses bureaux, cela peut créer une variation de distance domicile-travail à tous ses employés, indépendamment de leurs observables (excepté l'appartenance à l'entreprise). Les déménagements d'entreprise n'étant normalement pas présents dans les données, les auteurs contrôlent usuellement des changements d'entreprise, ou des changements de domicile (ou les sortent du champ de la régression), voire des changements de mode de transport lorsqu'ils sont disponibles et que le temps de transport domicile-travail est étudié. Cela semble garantir une inférence causale propre.

J'applique donc dans ce chapitre ces techniques à mon jeu de données, mais les complète aussi par une analyse descriptive de ce qui peut causer des variations dans la variable focale d'intérêt, le temps de trajet domicile-travail. M'intéressant

---

uniquement aux changements importants dans le temps de trajet domicile-travail qui ne sont pas causés par un changement d'entreprise ou de domicile, j'obtiens notablement que de l'ordre de 5% seulement des changements important de temps de trajet domicile-travail semblent causés par un déménagement de l'entreprise! D'autres explications qui échappent aux contrôles usuels sont possiblement des changements de poste (et de bâtiment) au sein d'une même entreprise, des variations régulières dans la distance domicile-travail pour certaines professions, des changements d'infrastructures modifiant le temps de déplacement domicile-travail, etc. Bénéficiant de la richesse du questionnaire de l'enquête Conditions de Travail, je propose un certain nombre de contrôles qui sont corrélés à une augmentation du temps de trajet domicile-travail et arrive ainsi à proposer une explication vraisemblable pour 50% des changements de trajet domicile-travail évoqués. Autrement dit toutefois, pour 50% des grands changements de trajet domicile-travail sans changement de domicile ni d'employeur, je ne sais pas ce qui cause cette variation. Je ne prétend donc pas que l'inférence, quoique améliorée par rapport à l'application des méthodes standards, bénéficie de la rigueur des méthodes quasi-expérimentales. Mais il me paraît difficile de traiter actuellement mieux la question posée avec les données d'enquêtes actuellement disponibles.

\* \* \*

## Recommandations de politiques publiques

**Le premier chapitre** conclut à des résultats ambivalents de la mise en place d'un jour de carence au l'indemnisation des arrêts maladies : la politique n'a pas réussi à réduire l'absentéisme des fonctionnaires de la Fonction publique de l'État, en ce que **la prévalence des absences pour raison de santé n'est pas significativement affectée. Leur distribution par durée l'est en revanche, avec une diminution de la prévalence des absences de courte durée** (d'exactement 2 jours), pouvant s'expliquer par un effet dissuasif de la pénalité financière

introduite par la mesure, **et une augmentation de la prévalence d'absences plus longues** (entre 1 semaine et 3 semaines). Trois mécanismes sont présentés pouvant expliquer cette hausse des longues absences. Tout d'abord, le jour de carence engendre un coût fixe pour le salarié à chaque prise d'arrêt maladie. Un agent n'a donc pas intérêt à hâter son retour au travail avant d'avoir la certitude d'être guéri. Ainsi, il peut trouver prudent de prolonger son arrêt, pour éviter une rechute synonyme d'une nouvelle pénalité. Ensuite, du fait de ce coût fixe certains agents connaissant un problème de santé pourraient hésiter à s'arrêter de travailler pour se soigner. Leur état de santé se dégraderait et conduirait *in fine* à des arrêts plus longs. Enfin, la mise en place d'un jour de carence pourrait générer chez des agents prenant un arrêt maladie le sentiment d'être injustement mis à contribution, les conduisant, par réaction, à prolonger un peu cet arrêt. Ces hypothèses ne sont ni exclusives ni exhaustives et aucun élément ne permet à ce stade dans l'article de les confirmer ou infirmer. Il me semble que plusieurs conséquences ou recommandations peuvent en être tirées.

D'une part, **si l'objectif de la politique est de diminuer les absences, elle a manifestement échoué.**<sup>6</sup> Même si nous ne comprenons pas ce qui cause l'augmentation des arrêts longs, notre étude et celles de Johansson et Palme (2005) ou de Davezies et Toulemon (2015), montrent qu'un système moins généreux peut ne pas causer significativement moins de journées d'absences, voire même peut sous certaines conditions causer plus de journées d'absences. A défaut d'une bonne compréhension théorique de ce qui est à l'oeuvre, une recommandation *a priori* raisonnable serait alors **d'adopter plutôt un changement déjà mis en oeuvre ailleurs, et qui a fait preuve d'efficacité** pour diminuer les absences. Ce pourrait être par exemple une diminution uniforme du taux de remplacement de 100% à 80%, comme cela a été le cas en Allemagne entre 1996 et 1999, qui a significa-

---

6. Il convient de noter que nos résultats ne sont pas incompatibles avec une diminution des arrêts maladies *stricto sensu*, ne tenant pas compte d'effets de reports vers d'autres causes. Pour ne donner qu'un exemple d'effet de report possible, dans un très récent document de travail portant sur une réforme du régime indemnitaire des arrêts maladies des fonctionnaires espagnols, les auteurs mettent en évidence des reports importants vers les accidents du travail, possiblement pour des douleurs de bas du dos (Marie et Vall Castelló, 2020).

---

tivement diminué les absences (voir par exemple Ziebarth et Karlsson (2014)<sup>7</sup>). Pour des motifs d'acceptabilité, cette diminution uniforme pourrait être plutôt de 90%, faisant notamment écho à l'abattement fiscal de 10% dont bénéficient tous les contribuables (y compris les employés du secteur public) pour tenir compte de leurs frais professionnels (frais kilométriques et repas), frais qu'ils n'encourent pas lors d'un congé maladie.

D'autre part, nous mettons en évidence que **la mesure a des effets hétérogènes** : parmi les variables observables que nous avons étudiées, la diminution des absences de 2 jours est plus importante pour les femmes, les jeunes employés et ceux qui travaillent moins de jours par semaine. On peut penser que d'autres variables non présentes dans les données pourraient avoir des impacts au moins aussi importants, tels que le fait d'occuper des postes difficilement remplaçables (Azmat, Hensvik et Rosenqvist, 2020), le temps de trajet domicile-travail... Similairement, les conséquences de la mesure, à savoir moins d'absences de très court terme et plus d'absences de long terme, ont probablement elles même des effets en terme de productivité qui varient probablement beaucoup en fonction des secteurs et des postes. Il est sûrement plus efficace de remplacer toute une semaine un agent relativement interchangeable, et membre d'une "chaîne de production", que d'avoir à le remplacer souvent quelques jours ; et au contraire de ne pas avoir à remplacer un agent autonome et aux compétences uniques, parce que cette personne n'aurait pas d'absences longues mais uniquement des absences courtes ne perturbant pas outre mesure son travail. Ceci interroge sur la pertinence d'un seul régime indemnitaire pour les arrêts maladies à toute la fonction publique, malgré la très grande diversité des emplois qu'elle recouvre. Après tout, le secteur privé, laissé au jeu des négociations paritaires, n'a pas convergé vers un seul système,

---

7. La seule réserve notée dans la littérature économique abondante sur cette expérience naturelle se trouve dans Pichler et Ziebarth (2017) : les auteurs constatent que les arrêts pour motifs non-contagieux diminuent significativement, mais pas ceux pour motifs contagieux ; ce que les auteurs interprètent comme une preuve indirecte de présentisme de personnes malades, favorisant les contagions. Pour limiter ce risque de contagion, les médecins pourraient avoir à indiquer dans le formulaire de l'arrêt maladie si le motif de l'arrêt implique, ou non, des risques de contagion au travail. Auquel cas la décote ne s'appliquerait pas.

mais de très nombreuses conventions collectives et accords d'entreprise.<sup>8</sup> Certaines dispositions sont peut-être optimales pour des métiers et des secteurs, mais pas à d'autres, comme suggéré par Lanfranchi et Treble (2010) sur la base de données françaises. **Avec un unique régime monolithique appliqué à tous ses agents, l'État français n'atteint peut-être pas une structure indemnitaire optimale** (du point de vue de l'employeur).

Enfin, **ces résultats doivent être mis en regard des développements liés à l'actualité sanitaire récente**. En effet, en mars 2020 au pic de la première vague de la covid-19, **le gouvernement français a supprimé le temps de la crise sanitaire les périodes de carence pour arrêt maladie** dans tous les secteurs, ceci «*dans le cadre de la gestion de l'épidémie et afin d'assurer une égalité de traitement de l'ensemble des assurés (mis en isolement, contraints de garder leurs enfants ou malades)*» (Gouvernement, 2020). Mais il a tardé à le faire au début de la seconde vague à l'automne 2020, déclenchant une tribune médiatisée dans Le Monde (Bréda et Toulemon, 2020). Leur argument principal est le suivant : citant notre étude (et d'autres études sur données françaises, celles de Davezies et Toulemon (2015) et de Ménard et Pollak (2015), un jour de carence semble augmenter le présentéisme. Ce qui pourrait favoriser la propagation des épidémies. Cette argumentation est raisonnable, et sans remettre en cause le bien-fondé de la tribune,<sup>9</sup> on notera toutefois que nous ne prouvons pas formellement une hausse du présentéisme du fait de la mesure, ni **ne mentionnons nous-même le risque de transmission d'épidémie comme facteur d'explication des absences d'une à trois semaines**. Ayant en effet consulté les durées de prescriptions d'arrêt maladie typiquement recommandées par l'Assurance Maladie suite à des avis de la Haute Autorité de la Santé (Assurance Maladie, 2020), celles-ci sont généralement de l'ordre de 3 à 5 jours dépendant des maladies infectieuses.<sup>10</sup> Sans

---

8. Tels que les trois jours de carence de la sécurité sociale sont finalement pris en charge pour deux tiers des employés du secteur privé (Perronnin, Pierre et Rochereau, 2012).

9. On notera en effet que d'autres études permettent d'étayer plus directement ces deux points : Stearns et White (2018), Pichler et Ziebarth (2017) et Marie et Vall Castelló (2020).

10. Angine : 3 jours ; Bronchite aiguë de l'adulte : 4-5 jours pour un travail sédentaire ou un travail physique léger, sinon 7 jours ; gastro-entérite virale : 3 jours ; grippe saisonnière : 5 jours ; sinusite maxillaire : 3 jours.

---

l'exclure définitivement, cette hypothèse ne nous semble donc pas expliquer les augmentations significatives constatées. **Peut-être cela pourrait-il en revanche expliquer pourquoi nous n'observons pas de diminution significative des absences comprises entre 3 jours et 1 semaine** : une moindre incidence liée à des pénalités financières serait fortement compensée par une hausse des maladies infectieuses, ainsi que par le ou les autres phénomènes expliquant les hausses des absences d'une à trois semaines.

Le deuxième chapitre a également plusieurs enseignements principaux nuancés :

Tout d'abord, nous avons montré que, à elle seule, **la politique du tarif unique seule n'a généralement pas eu d'impact significatif sur le taux de chirurgie ambulatoire des GHMs étudiés dans le secteur privé**, ni dans les premières vagues d'incitations (à partir de 2008 ou de 2009), ni dans la vague ultérieure étudiée (à partir de 2012). Elle a pu avoir un impact significatif sur certains GHMs lorsque le tarif unique était accompagné de la deuxième politique, la mise sous accord préalable (MSAP)<sup>11</sup>, mais ce lien n'est pas systématique. En revanche, **le tarif unique a eu un impact significatif la plupart du temps dans le secteur public pour les GHMs incités en 2008 ou 2009**, quel que soit le niveau de la MSAP. En cela, nos résultats permettent de compléter les conclusions de Allen, Fichera et Sutton (2016) : *«Les payeurs peuvent agir proactivement dans la fixation des tarifs et peuvent» parfois «s'attendre à une réaction importante des établissements de santé».*

De plus, **les taux de chirurgie ambulatoire de certains GHMs qui bénéficiaient depuis 2008 ou 2009 de la politique de MSAP (dans un sens lâche décrit dans l'article) ont significativement augmenté, en particu-**

---

11. Cette politique consiste en ce que l'Assurance Maladie peut imposer à un hôpital, pour un ensemble d'actes médicaux définis et pour une période limitée dans le temps, d'avoir à justifier préalablement toute opération ne s'effectuant pas en chirurgie ambulatoire. Si la justification n'est pas établie aux yeux de l'Assurance maladie, celle-ci peut refuser le paiement de l'opération réalisée en chirurgie conventionnelle. Dans l'article, nous parlons de MSAP lâche pour désigner les groupes d'actes pour lesquels l'Assurance Maladie peut appliquer ces mesures, sans distinguer les hôpitaux qui les ont effectivement subi.



**lier dans le secteur public.** Des effets significatifs et persistants s'observent également pour des GHMs dont un nombre moyennement important de séjours étaient susceptibles d'être concernés par la MSAP, ce qui laisse penser que l'effet de la MSAP sur les seuls séjours potentiellement concernés est assez important. **Nous l'interprétons par exemple comme un « effet d'apprentissage » de la MSAP,** car les effets estimés sont nettement plus forts que ceux qui auraient été obtenus si seuls les séjours effectivement placés en MSAP (au sens strict) avaient été affectés. En effet, il est plausible que des services aient immédiatement réagi à la publication à un niveau national de listes d'actes qui peuvent être souvent réalisés en chirurgie ambulatoire et qui sont susceptibles d'être placés en MSAP, si l'hôpital a un faible taux d'ambulatoire pour ces actes.

Par ailleurs, **la rareté des réactions significatives pour les GHMs incités à partir de 2012** n'entre pas en contradiction avec cette hypothèse d'« effet d'apprentissage », voire la renforce. Par exemple, lorsque la politique de mise sous accord préalable est passée de 129 actes en 2009 à 200 actes en 2012, il est possible que les acteurs aient été moins convaincus de la pertinence de la pratique ambulatoire sur les actes rajoutés (puisqu'ils n'avaient pas été directement intégrés en 2008/2009). Il a aussi pu y avoir des effets d'apprentissage d'un autre type, liés ici à l'application de la politique de MSAP : les acteurs hospitaliers ont pu réaliser que les placements en MSAP restaient relativement « rares » par rapport aux possibilités très étendues offertes par la loi à l'Assurance Maladie.<sup>12</sup> **Par conséquent, une interprétation de nos résultats est que la politique de surveillance peut avoir des effets importants, mais uniquement si elle est crédible :** soit parce que les actes visés peuvent effectivement être facilement appliqués en chirurgie ambulatoire de manière standard, soit parce que les hôpitaux peu performants au regard de la pratique de certains actes visés sont effectivement l'objet de MSAP (au sens strict).

En outre, parmi les groupes dans le secteur public dont le taux de chirurgie am-

---

12. En effet, cette dernière peut en théorie placer l'établissement en MSAP s'il est en dessous de la moyenne régionale ou nationale pour cet acte ; or, même pour des GHMs avec de très forts taux de MSAP au sens lâche, nous n'avons observé que peu de MSAP au sens strict les premières années d'incitation.

---

bulatoire a crû du fait de la politique de MSAP, **la mesure ne semble généralement pas avoir augmenté le nombre total de séjours ni conduit à une augmentation du nombre de réadmissions**. Même s'il serait préférable d'avoir des indicateurs reflétant directement la santé et le bien-être des patients, ce dernier point suggère que les augmentations de chirurgie ambulatoire suscitées par la mesure ne l'ont pas été au détriment de la santé des patients.

Enfin, **le secteur public semble avoir réagi plus souvent aux deux politiques étudiées**. La différence de résultat entre les deux secteurs pourrait venir de ce que la marge d'augmentation de la pratique de la chirurgie ambulatoire ait été plus grande dans le secteur public que dans le secteur privé.

Il me semble que plusieurs enseignements peuvent être tirés de ces résultats, et de ce deuxième chapitre.

**Premièrement** et de façon triviale, le rare impact du tarif unique dans le secteur privé, peut-être parce que ce secteur s'était déjà engagé de façon autonome dans une augmentation de la chirurgie ambulatoire, rappelle que **la mise en place d'une incitation financière ne suffit pas à produire un effet**. Il est nécessaire qu'un effet soit possible, par exemple que des marges d'amélioration existent. A l'instar de ce qui a été réalisé par Lefebvre-Hoang et Yilmaz (2019), ceci pourrait (par exemple) indirectement être observé en amont de la mise en place de la politique *via* des disparités de pratique de la chirurgie ambulatoire entre secteurs, entre zones géographiques, entre établissements (et en contrôlant des caractéristiques de la patientèle et la composition des séjours des hôpitaux, voir Dormont et Milcent (2013)).

**Deuxièmement, les effets observés pour les deux politiques étaient loin d'aller de soi**. C'est la direction de l'hôpital, consciente des enjeux afférents aux recettes de l'hôpital, qui est visée par le mécanisme de tarif unique, qui ne peut vraisemblablement toucher les médecins qu'au travers de la communication de la direction de l'hôpital. Il n'est pas évident qu'une telle incitation serait efficace du fait des problèmes « principal-agent » entre la direction de l'hôpital et ses méde-

cins. **Les effets significatifs que nous obtenons suite à la mise en place d'un tarif unique**, notamment dans le secteur public pour les GHMs incités à partir de 2008 ou de 2009, **conduisent à penser que les directions d'hôpital ont réussi** dans ce cas de figure à entraîner leurs médecins dans une augmentation de la pratique ambulatoire pour les GHMs ciblées par cette incitation financière.

**A *contratio*, la MSAP implique à la fois la direction et les médecins.**

La direction est concernée par le mécanisme de MSAP, puisqu'en cas de refus de l'Assurance Maladie, c'est aussi le revenu de l'établissement qui est réduit. Mais la MSAP implique surtout les médecins eux-mêmes, puisqu'en cas de surveillance par l'Assurance Maladie, ils doivent justifier leurs décisions et suivre une procédure administrative contradictoire (avec un autre médecin). **Cet alignement de tous les acteurs pour développer la pratique ambulatoire des actes ciblés peut notamment expliquer pourquoi nous obtenons des effets significatifs** pour les GHM impliquant des actes placés en MSAP, sans commune mesure avec ce qu'on attendrait si les acteurs réagissaient aux seuls actes pour lesquels leur établissement était placé en MSAP. Ces effets peuvent éventuellement venir d'un « effet d'apprentissage », à savoir la déclaration officielle que certains actes peuvent être souvent pratiqués en ambulatoire, ou peut-être d'un « effet de précaution », lié à la crainte de subir un fardeau administratif supplémentaire impliquant pour les médecins une procédure contradictoire de leurs décisions et d'éventuels refus, coûteux pour l'administration de l'hôpital, si le service est repéré comme pratiquant moins la chirurgie ambulatoire que les autres services alentours.

**Troisièmement, il me semble que plus de cohérence entre grands acteurs institutionnels auraient probablement été bénéfiques.** La politique du virage ambulatoire a certes bénéficié sur la durée d'une assez grande cohérence temporelle, avec l'augmentation du champ des deux politiques étudiées, un suivi régulier par les tutelles, la mise à disposition croissante d'outils aux établissements. Mais les deux politiques du tarif unique et de la MSAP étaient présentées suivant des grilles différentes, ne se recouvrant que partiellement ; l'une sur des actes pouvant être pratiqués en chirurgie ambulatoire de façon usuelle, l'autre sur

---

des GHMs sélectionnés de façon non-publique (potentiellement avec des visées budgétaires). Pour faciliter la visibilité des deux dispositifs, l'État et l'Assurance Maladie auraient pu, au moins dans un premier temps, se coordonner ouvertement pour que le tarif unique soit présent pour les GHMs fortement concernés par la MSAP. Selon notre hypothèse d'effets d'apprentissage, une telle politique aurait pu avoir des effets peut-être plus importants encore, en mobilisant très fortement tous les acteurs autour d'une seule politique motivée et simple.<sup>13</sup>

Enfin, il est assez clair que la perspective de l'évaluation n'a pas été prise en compte au moment du choix de la politique. Cela pourrait ressembler à un voeu pieux, mais des possibilités existent. Par exemple, plutôt que d'augmenter le champ de la politique de façon incrémentale jusqu'à sa généralisation en seulement six ans, un premier ensemble de GHMs ayant vocation à être concerné à terme aurait pu être identifié, et seulement une moitié, tirée au hasard, aurait pu être concernée pendant les trois premières années.

*A contrario*, les résultats du troisième chapitre sont plus directs. Bien que j'obtienne en coupe que le temps de trajet entre le domicile et le travail augmente significativement arrêts maladies et présentéisme des employés, **ces effets ne sont plus statistiquement significatifs lorsque les caractéristiques individuelles fixes dans le temps sont prises en compte**. En panel, et à l'instar de la littérature, j'obtiens en revanche que **le temps de trajet domicile-travail peut affecter négativement la santé auto-évaluée et le bien-être mental des femmes**.

Il me semble que ces résultats intéresseront particulièrement au regard **du choc exogène que l'épidémie de covid-19 risque de créer sur le développement du télé-travail**. Pour des études ultérieures sur son impact sur les absences et le présentéisme, **il ne semble pas que l'arrêt (ou la diminution) du déplacement domicile-travail soit en lui-même de nature à influencer significativement les absences et le présentéisme**. En revanche, **cette diminu-**

---

13. Bien évidemment, cela réduirait à néant la possibilité d'évaluer distinctement les deux politiques.

tion du déplacement domicile-travail pourrait influencer positivement l'état de santé, et notamment le bien-être psychologique, notamment des femmes.<sup>14</sup>

---

14. Et peut-être aussi des hommes impliqués à parité dans les tâches du foyer, si cela est l'une des clés expliquant l'impact genré que j'obtiens.

## Bibliographie

- Abadie, Alberto, Alexis Diamond, et Jens Hainmueller.** 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American statistical Association*, 105(490): 493–505.
- Allen, Thomas, Eleonora Fichera, et Matt Sutton.** 2016. "Can payers use prices to improve quality? Evidence from English hospitals." *Health economics*, 25(1): 56–70.
- Azmat, Ghazala, Lena Hensvik, et Olof Rosenqvist.** 2020. "Workplace presenteeism, job substitutability and gender inequality."
- Bozio, Antoine.** 2014. "L'évaluation des politiques publiques: enjeux, méthodes et institutions." *Revue française d'économie*, 29(4): 59–85.
- Bozio, Antoine, Pierre Cahuc, Philippe Choné, Anne Epaulard, Marc Ferracci, Giuseppe Nicoletti, Anne Perrot, Jérôme Philippe, André Sapir, et Etienne Wasmer.** 2015. "Conclusions de la Commission d'étude des effets de la loi pour la croissance et l'activité." Commission d'étude des effets de la loi pour la croissance et l'activité Report.
- Bréda, Thomas, et Léa Toulemon.** 2020. "Covid-19 : "Pour limiter l'épidémie", en incitant à "rester chez soi au moindre symptôme, le gouvernement devrait prendre en charge les jours de carence"."
- CAE.** 2013. "Evaluation des politiques publiques." *Notes du conseil d'analyse économique*, , (1): 1–12.
- Davezies, Laurent, et Léa Toulemon.** 2015. "Does moving to a system with a more generous public health insurance increase medical care consumption?" *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, , (119-120): 179–205.

- Dormont, Brigitte, et Carine Milcent.** 2013. “Comment évaluer la productivité et l’efficacité des hôpitaux publics et privés? Les enjeux de la convergence tarifaire.”
- Givord, Pauline.** 2014. “Méthodes économétriques pour l’évaluation de politiques publiques.” *Economie prevision*, , (1): 1–28.
- Gouvernement.** 2020. “Amendement no204 (Rectificatif) à la Loi d’urgence pour faire face à l’épidémie de covid-19.”
- Heckman, James J.** 2020. “Randomization and Social Policy Evaluation Revisited.”
- Johansson, Per, et Mårten Palme.** 2005. “Moral hazard and sickness insurance.” *Journal of Public Economics*, 89(9): 1879–1890.
- Lanfranchi, Joseph, et John Treble.** 2010. “JUST-IN-TIME PRODUCTION, WORK ORGANIZATION AND ABSENCE CONTROL.” *The Manchester School*, 78(5): 460–483.
- Lefebvre-Hoang, Ingrid, et Engin Yilmaz.** 2019. “État des lieux des pratiques de chirurgie ambulatoire en 2016.”
- maladie, Assurance.** 2020. “Mémorandum de référentiels de durée pour les maladies infectieuses.”
- Marie, Olivier, et Judit Vall Castelló.** 2020. “If Sick-Leave becomes More Costly, Will I go back to Work? Could it be too soon?”
- Ménard, Samuel, et Catherine Pollak.** 2015. “L’effet d’une extension des indemnités complémentaires sur les arrêts maladie. Une évaluation de l’ANI de 2008.” *Dossier Solidarité et Santé*, 69.
- Perronnin, M., A. Pierre, et T. Rochereau.** 2012. “L’enquête Protection sociale complémentaire d’entreprise 2009.” *Les rapports de l’Irdes*, , (554).

- Pichler, Stefan, et Nicolas R Ziebarth.** 2017. “The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior.” *Journal of Public Economics*, 156: 14–33.
- Rubin, Donald B.** 1974. “Estimating causal effects of treatments in randomized and nonrandomized studies.” *Journal of educational Psychology*, 66(5): 688.
- Stearns, Jenna, et Corey White.** 2018. “Can paid sick leave mandates reduce leave-taking?” *Labour Economics*, 51: 227–246.
- Ziebarth, Nicolas R, et Martin Karlsson.** 2014. “The effects of expanding the generosity of the statutory sickness insurance system.” *Journal of Applied Econometrics*, 29(2): 208–230.





# Chapitre 1

Les effets d'un jour de carence pour  
arrêt maladie sur les absences pour  
raison de santé dans la fonction  
publique de l'État française

**Résumé :**

Cet article étudie les effets d'un jour de carence pour les congés maladie. Il a été mis en place dans la fonction publique française en 2012-2013. Nous employons une stratégie de différence de différences avec effets fixes individuels, en utilisant les données de l'Enquête Emploi en Continu. Nous constatons que la prévalence totale des absences pour raison de santé n'est pas affectée par cette politique moins généreuse. En revanche, la distribution de leur durée l'est. La prévalence des absences de courte durée diminue, tandis que la prévalence des absences de longue durée augmente. La diminution des absences de courte durée est plus importante pour les femmes, les jeunes et ceux qui travaillent peu de jours par semaine. Dans l'ensemble, nous concluons que la politique n'a pas réussi à réduire l'absentéisme.

Ce chapitre a été co-écrit avec Alexandre Godzinski (Insee, Crest & PSE).

\* \* \*

---

Lower sick leave cover,  
fewer health-related work absences? <sup>a</sup>

Alexandre Cazenave-Lacroutz <sup>b, c, d</sup>      Alexandre Godzinski <sup>b, e</sup>

**Abstract**

*This paper studies the effects of a one-day waiting period for sick leave. It was implemented in the French civil service in 2012-2013. We employ a difference-in-differences strategy with individual fixed effects, using Labour Force Survey data. We find that the total prevalence of health-related absences is not affected by this less generous policy. However its duration distribution is. The prevalence of short-term absences decreases, while the prevalence of long-term absences increases. The decrease in short-term absences is higher for women, young employees and those working fewer days per week. Overall, we conclude that the policy failed to reduce absenteeism.*

**JEL:** D82, I18, J22, J45.

**Keywords:** Absenteeism, Sickness pay, Public sector, Monetary incentives.

---

<sup>a</sup>This work was previously entitled "Effects of the one-day waiting period for sick leave on health-related absences in the French central civil service". We thank Didier Blanchet, Pierre Cahuc, Eve Caroli, Arthur Charpentier, Andrew Clark, Brigitte Dormont, Philippe Février, Pierre-Yves Geoffard, Fanny Godet, Malik Koubi, Joseph Lanfranchi, Pierre Picard, Sébastien Roux, and Joachim Winter as well as seminar audience at the 2018 EEA-ESEM conference in Cologne, the 2018 EALE conference in Lyon, the 2018 EuHEA conference in Maastricht, at Crest, Insee and Paris-Est University for helpful comments. Declarations of interest: none.

<sup>b</sup>Insee-Crest, 88 Avenue Verdier, 92541 Montrouge Cedex, France

<sup>c</sup>Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris

<sup>d</sup>Corresponding author: alexandre.cazenave-lacroutz@polytechnique.org

<sup>e</sup>Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris.

# 1 Introduction

---

The question of how health-related absenteeism reacts to the generosity of reimbursement patterns remains an empirical concern, due to the social costs of absence from work. In a simple theoretical framework with unidimensional effort choice and unidimensional coverage level, the classic implication is that a lower coverage leads to a higher effort. The effort choice related to sick leave is however bidimensional, as the worker may have some latitude regarding both whether she starts a sick leave and how long this leave lasts. The coverage is itself multidimensional, since the replacement rate may vary over the days of sick leave. In accordance with the unidimensional model, several papers on sick pay reforms across Europe find that the prevalence of absence decreases when the generosity of sick pay decreases (Ziebarth and Karlsson, 2014; Chemin and Wasmer, 2009; Henrekson and Persson, 2004). But subtler results have recently been found in settings that are more distant from the unidimensional reimbursement framework (Davezies and Toulemon, 2015; Paola, Scoppa and Pupo, 2014; Pettersson-Lidbom and Thoursie, 2013; Johansson and Palme, 2005). Studying the implementation of a waiting period or similar measures, most papers find that a lower generosity during the first days of sick leave induces a decrease in the number of short-term sick leave. But it can also induce an increase in the duration of long-term sick leaves. As a result, it does not necessarily lead to a decrease in total prevalence. For instance, after the abolishment of a waiting period, it was sometimes found that total prevalence had significantly increased (Pertold, 2019), but sometimes not (Davezies and Toulemon, 2015) and sometimes a significant decrease in total prevalence was even found (Pettersson-Lidbom and Thoursie, 2013).<sup>1</sup> Consequently, the reaction of health-related absenteeism to a change in the generosity of the reimbursement pattern remains an open research field.

This paper evaluates the effects of the presence of a one-day waiting period for sick leave on the prevalence of health-related absences, defined as the proportion of employees absent from work for health-related reasons. We also differentiate the effects between different categories of sick leave spells, constructed from information on durations. For that purpose, we exploit two exogenous changes in sick leave pay in the French civil service. On 1<sup>st</sup> January 2012, the French government introduced a one-day waiting period for all

---

<sup>1</sup>See Pollak (2017) for a review of this literature.

workers in the French civil service to combat absenteeism. The first day of sick leave was hence no longer paid. On 1<sup>st</sup> January 2014, exactly 2 years later, the following government repealed the measure. Both the exogenous introduction of this one-day waiting period and its exogenous repeal create an ideal quasi-natural experiment to assess this component of a sick pay scheme.

We apply a strategy of difference-in-differences between the employees of the French central civil service and those of the private sector. We choose to focus on the central civil service, the largest of the three parts of the French civil service, for three reasons. First, in the hospital and territorial civil services, the two other parts of the French civil service, other monetary incentives related to work attendance exist. The characteristics and the timing of implementation of these other incentives differ greatly between public institutions and over time. Second, in the territorial civil service, the timing of implementation of the one-day waiting period also varied greatly between public institutions and over time. Third, regarding sick leave trends, the hospital civil service is not comparable with the private sector as a whole or with the private hospital sector.

We use a rotating panel: the French Labour Force Survey, from 2010 to 2014. We can thus include individual fixed effects to control for unobserved individual heterogeneity.

We obtain three different results. First, we do not find that the one-day waiting period decreases health-related absences. If anything, it increases them. Second, it does however lead to a change in the duration distribution of these absences. Short-term absenteeism decreases, while long-term absenteeism increases. More precisely, we find that there is a significant decrease of 60% in the prevalence of 2-day sick leave spells, and a significant increase of 50% in the prevalence of 1- to 3-week sick leave spells. Third, effects differ across sociodemographic characteristics and working conditions. The decrease in short-term absences is higher for women, young employees and those working only a few days per week.

Our study has four main advantages over previous papers focusing on day waiting periods or similar schemes. First, we study the impact of a change in one single parameter of the sick pay pattern, the replacement rate of the first day of sick leave, whereas for

instance Pettersson-Lidbom and Thoursie (2013) get a mixed effect of the abolishment of a one-day waiting period and of an increase in the replacement rate for sick leave up to 14 days. This enables us to isolate the specific impact of a one-day waiting period on health-related absences. Second, the introduction and the repeal of the measure allows us to assess the symmetry of the corresponding effects and to conduct robustness checks. Third, as we use difference-in-differences on large groups, we are not exposed to the suspicion that some local sickness shock may exacerbate our results, as highlighted by D'Amuri (2017) regarding some of Paola, Scoppa and Pupo (2014) 's results. Fourth, works that use survey data in the related literature rarely control for unobserved individual heterogeneity, which we do through the inclusion of individual fixed effects. Besides, this is the first research paper that assesses the effects of this controversial policy change in France.

The remainder of the paper is organized as follows. In Section 2, we present the institutional framework, the data and the descriptive statistics. In Section 3, we present the empirical strategy. Our results regarding the responses of the prevalence of health-related absence and the prevalence by duration categories are presented in Section 4, as well as those regarding the heterogeneity of these responses. Some robustness tests follow in Section 5. Section 6 concludes.

## 2 Institutional framework, data, and descriptive statistics

### 2.1 The 1<sup>st</sup> January 2012 introduction and the 1<sup>st</sup> January 2014 repeal of the one-day waiting period in the French civil service

Until the end of 2011, employees of the civil service benefited from a 100% replacement rate of their wage for the first 3 months of their sick leave.<sup>2</sup> After that threshold, the replacement rate fell to 50% of their wages (unless they had taken out optional additional coverage). Hence they enjoyed full coverage for sick leave before that threshold, and partial coverage thereafter (unless optional full coverage). In both cases, a medical certificate had to be provided to the employer.

In mid-November 2011, the French government announced that an amendment to the October draft of the 2012 Budget Act for 2012 would implement a one-day waiting period in the civil service. The measure was taken by the right-wing ruling party for reasons of equity with respect to the private sector and also to reduce absenteeism. This monetary incentive was strong since it caused the replacement rate to fall on the first day from 100% to 0%.<sup>3</sup> The measure was effective on the 1<sup>st</sup> January 2012. The date of effect hence followed the date of announcement fairly closely, limiting possible anticipation effects. This policy applied to the whole civil service, that is, all civil servants, soldiers, and employees under a private contract in the civil service.<sup>4</sup>

---

<sup>2</sup>This applied to all civil servants in the central civil service, and to all employees under a private contract in the central civil service and whose seniority was above 3 years. Both basic salary and bonuses were subject to a 100% replacement rate for 3 months. For employees under a private contract in the civil service and whose seniority was below 3 years, the period with no wage loss was at most 2 months. For employees of the civil service outside the central civil service, the same rules applied, except that only the replacement rate of the basic salary was defined by the law. The replacement rate of bonuses may follow different rules.

<sup>3</sup>Technically, the wage penalty was equal to 1/30 of the usual monthly wage, whatever the calendar month and the number of working days per week.

<sup>4</sup>There was a doubt as to whether previously state-owned companies that still had many civil servants had to apply it, but it appeared that this was not necessarily the case after the French public transport company in Paris and suburbs (RATP) was successfully sued because it had started applying it. Other similar firms, such as the French Postal Service, considered they were not required to apply it to their civil servants.



Implementation details were specified by a circular dated 24<sup>th</sup> February 2012. The policy concerned neither work accident leave nor the so called "long duration" and "long sickness" leaves (both cover serious diseases such as cancers), and neither maternity leave nor parental leave. Since the implementing circular was signed only in February 2012, and since many difficulties in adapting the pay information systems to the policy were reported, it is likely that the first deductions of earnings started with some delay. However, the circular clearly stated that it applied to all sick leaves starting from 1<sup>st</sup> January 2012 and the measure was highly publicized (notably by labour unions). Hence most employees in the central civil service and in the hospital civil service had heard of the change, and probably knew it was to be enforced as soon as 1<sup>st</sup> January 2012. The territorial civil service differed in that there was considerable heterogeneity in the policy implementation.

The possibility of covering the one-day waiting period by a collective health insurance plan also differed between, the central and hospital civil services on the one hand, and the territorial civil service on the other. While coverage may have been available in some units of the territorial civil service, we are unaware of such coverage in the central and hospital civil services.

During the presidential campaign, the left-wing contender promised to abolish the one-day waiting period if elected. He became president in May 2012. In the 2014 Budget Act, the left-wing ruling party removed the one-day waiting period for sick leaves for all civil service employees.<sup>5</sup> The measure started on 1<sup>st</sup> January 2014 and there is no reason to believe it did not come into force immediately.

In the French private sector, the social security compensates sick leave by providing sick leave benefits equal to half of the wage after a three-day waiting period. After a seven-day waiting period, the employer is also obliged to contribute, so that benefits then reach at least 90% of the wage for the following 30 days. But most employees benefit from more generous conditions than those strictly required by the law. For instance, in 2009,

---

<sup>5</sup>A reinforcement of monitoring was announced at the same time. If the physician certificate was not sent within 48 hours after the sick leave started, civil servants could lose half of their benefit between the date of prescription and the date of transmission of the physician's certificate. Note that the corresponding decree was published in October 2014 and the corresponding circular was released in April 2015. It was also announced that controls of the relevance of sick leave would be increased.

two-thirds of private sector employees had full coverage for the first three days of their sick leave (Perronnin, Pierre and Rochereau, 2012). This situation is very heterogeneous since it is dependent upon collective agreements at the industry or employer level. A medical certificate is always required. Note that in July 2008, some of the sick leave rules were reformed. There was no other change regarding sick leave rules in the private sector between July 2008 and 2014.

The introduction and the repeal of the one-day waiting period policy constitute two quasi-natural experiments. They affected only the civil service, and did not concern the private sector. We choose to focus on the central civil service, the largest of the three parts of the French civil service, for three reasons. First, in the territorial and hospital civil services, other monetary incentives aiming at reducing absenteeism exist. They include for example semiannual or annual bonuses that may depend on work attendance. The characteristics and the timing of the implementation of these other incentives vary greatly between public institutions and over time. Second, in the territorial civil service, the timing of the implementation of the one-day waiting period also varied greatly between territorial authorities and over time.<sup>6</sup> Third, in terms of sick leave trends, the hospital civil service is not comparable with the private sector as a whole or with the private hospital sector.

## 2.2 Data: the French Labour Force Survey

This study uses the French Labour Force Survey. It is a rotating panel, with 100,000 individuals being interviewed for 6 subsequent quarters. Each quarterly interview focuses on the reference week, defined as the week just before the interview.<sup>7</sup> Responding to the survey is legally mandatory.

The Labour Force Survey is to our knowledge the only available dataset for France

---

<sup>6</sup>In an answer to an oral question in the French Senate published on 27<sup>th</sup> March 2013 and related to the non-application of the one-day waiting period in some territorial civil services, the French Minister of civil service stated that she could "*understand that [the one-day waiting period] would not necessarily be applied in the following months of its existence. Each local authority executive should decide what to do*" (translation). See Sénat (2013).

<sup>7</sup>This reference week is randomly sampled within the calendar quarter of the first interview. Each subsequent interview then occurs with a time interval of precisely one quarter.

that includes information relative to absences for sickness in both the civil service and the private sector over the period. It also allows us to properly focus on absences due to health reasons, whereas administrative data<sup>8</sup> focus on absences for which a medical certificate is provided. In both cases, these absences may be subject to prevention efforts to avoid getting sick, which corresponds to *ex ante* moral hazard, and to several layers of hidden actions after the employee knows she is sick, which corresponds to *ex post* moral hazard. These layers potentially include the decisions to go to work, to choose a doctor, to consult her and to influence her decision regarding the medical diagnosis and the medical certificate. But the difference is that there is an extra layer of *ex post* moral hazard with the administrative data, which is the reason given by the employee to her employer for her absence. For short-term absences, employees might use days off in order to avoid a wage penalty. Survey data are not sensitive to such reporting choices.

In the survey, two different sequences of questions can be used to determine if the survey respondent was absent from work for health reasons. The use of one or the other sequence of questions depends on whether the individual worked at least one hour during the reference week or not. In the former case, when the individual worked for at least one hour during the reference week, she is asked whether she was absent because of sickness or a work accident<sup>9</sup> and how many days during the reference week this absence lasted. In the latter case, when the individual did not work at all during the reference week, she is asked why she did not work. One of the possible answers is sick leave (including

---

<sup>8</sup>Examples of papers using administrative data include Pettersson-Lidbom and Thoursie (2013) and Paola, Scoppa and Pupo (2014).

<sup>9</sup>The regulation on sick leave and on leave related to a work accident are distinct. In particular, no waiting period for leave related to a work accident was implemented in 2012 in the civil service. This would tend to attenuate our estimates. Note however that over a reference day the prevalence of absence for work accident is ten times lower than the prevalence of absences for sick leave in the central civil service (DGAFP, 2015). In the private sector, a decree in 2010 changed the way contribution rates of companies are estimated and took effect from 2012 on. The official role of occupational doctors was also modified in July 2012. But this firstly applies to the employer side of the sick leave scheme. Secondly, expenditures are also more than ten times lower for work accidents than for sick leave in the private sector (Drees, 2014).

## *1.2. Institutional framework, data, and descriptive statistics*

leave for a sick child<sup>10</sup>) or leave related to a work accident. When this answer is chosen, the individual is then asked the expected total duration of the leave. More details and extracts from the questionnaire are available in Appendix A. For the sake of brevity, unless otherwise specified, we refer to these health-related absences as sick leave.

The two durations that correspond to the two sequences of questions have a different meaning. In the first case, it is a realized value, but the duration may be left-censored or right-censored, as the sick leave may have begun before or may continue after the reference week. In the second case, it is an expected value, but which is related to the total duration of the sick leave.

As a result of these two different intrinsic meanings of duration in the data, a duration model analysis cannot be conducted. To go beyond the mere study of total prevalence of sick leave, we break sick leave spells down into different categories that are based on the reported spell length. We then study the reaction of each category of sick leave spells to the policy. The stability of the definition of these categories across groups and time allows to get a better insight of the effect of the policy on the pattern of sick leave spells, even if no quantitative conclusion can be drawn on incidences or on durations from our results.

We refer to sick leave spells as "short-term spells" when the interviewed person worked at least one hour during the reference week (which corresponds to the first sequence of questions mentioned above). We call sick leave spells "long-term spells" when the interviewed person did not work at all during the reference week (which corresponds to the second sequence of questions mentioned above). By construction, the duration of the former cannot exceed 7 days, and the duration of the latter is rarely under 7 days.<sup>11</sup>

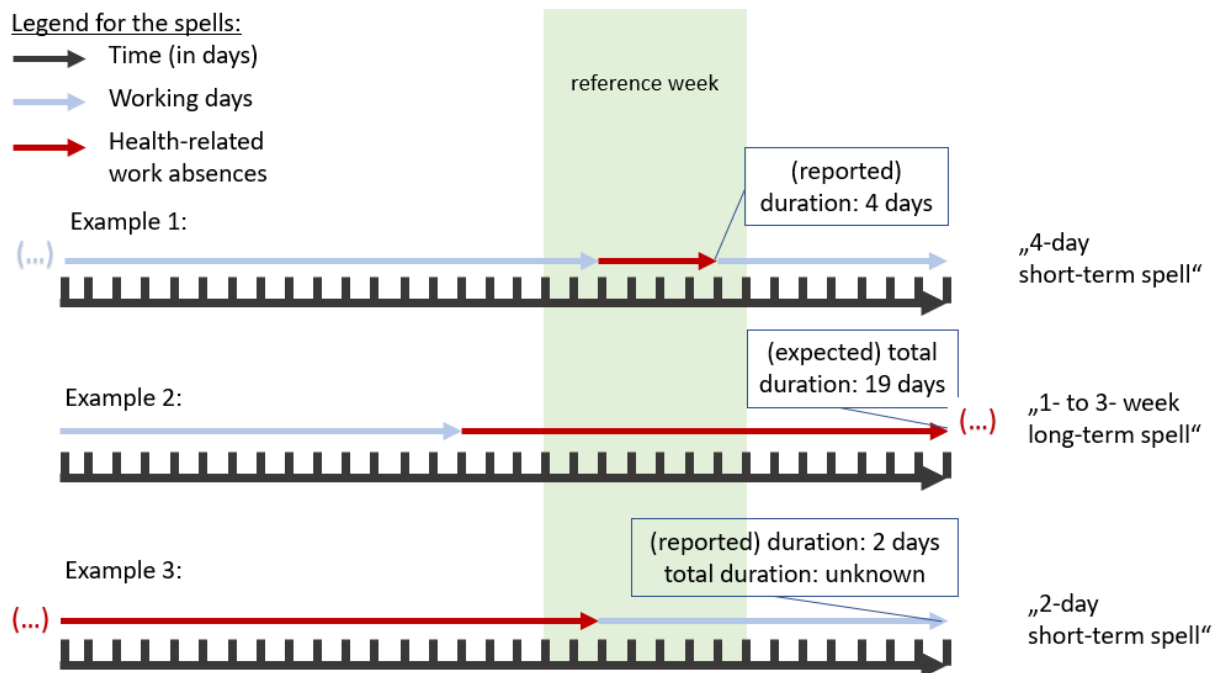
---

<sup>10</sup>The questionnaire specifies that sick days for care of a sick child should also be included in this second case. In the first case, the wording regarding whether they should be included is not explicit. Since no waiting period for sick days for care of a sick child was implemented in 2012 in the civil service, this might tend to attenuate our estimates since it mixes with absences related to the own health of the interviewed person. Nevertheless, over a reference day the prevalence of absences for a sick child is twenty-five times lower than the prevalence of absences for sick leave in the central civil service (DGAFP, 2015), so this is negligible. We also note that, in the presence of a one-day waiting period, some parents could possibly report leave for a sick child to their employer instead of sick leave, since the former has no one-day waiting period. However, in those cases, a document stating that the parent's presence was necessary is required.

<sup>11</sup>This can happen in case of part time jobs for instance: the person has been absent only two days, but these two days were her working days. This rarely occurs.

To break sick leave spells down into precise duration categories, we consider both kinds of spells with durations of 1 day, 2 days, and 3 to 7 days, and long-term spells with durations between 1 and 3 weeks, between 3 weeks and 3 months, and over 3 months<sup>12</sup>. The 1-week threshold was chosen due to the structure of the questionnaire. The 3-month threshold was chosen because employees of the civil service without additional coverage start losing half of their benefits after that threshold. The 3-week threshold was eventually chosen to further investigate preliminary results on 1-week to 3-month sick leave (Cazenave-Lacroutz and Godzinski, 2017). We have a breakdown of sick leave spells into 6 categories: each spell is in one and only one category. Figure 1 illustrates on a few examples how this partition works.

**FIGURE 1.**  
Timeline for selected examples of sick leave spells.



We define prevalence as the proportion of individuals who are on sick leave during the reference week. It differs from incidence, which is the proportion of individuals who begin a sick leave during that week. Both are of interest, but our data do not afford access to incidence, since we do not know when the sick leave begins. We thus focus on prevalence. We consider the prevalence of all sick leave spells but also the prevalence of each category of spells.

<sup>12</sup>As a consequence, the rare "long-term" spells under 1 week are grouped with "short-term" spells.

### *1.2. Institutional framework, data, and descriptive statistics*

For the descriptive statistics as well as the regressions, we use weights produced by the French Statistical Institute (INSEE). Due to the limited sample size of the central civil service, we cap the value of the highest weights (1%). Note that we use cross section weights since longitudinal weights are not currently available for the French Labour Force Survey. In order to use the weighted regression with individual fixed effects, we need a unique weight per individual. Thus, we attribute to each individual a unique weight equal to the mean of her weights over the periods of observation.

We limit ourselves to individuals aged between 15 and 75, due to the design of the French Labour Force Survey. In order for the private sector to be a convincing counterfactual to the central civil service, we only keep wage earners that are civil servants or under permanent contracts. Both categories benefit in France from strong employment protection that distinguish them from self-employed workers and employees under fixed-term contracts, and employment protection is known to be a key parameter in health-related absenteeism (Ichino and Riphahn, 2005). Moreover, civil servants stand for 90% of employees of the central civil service. We also exclude survey respondents without information on their wage or on whether they have been absent from work during the reference week, and those that were not exclusively in the central civil service or exclusively in the private sector.

## **2.3 Descriptive statistics**

We begin by looking at the total sick leave prevalence over time and by sector<sup>13</sup> (Figure 2), on both a yearly and a quarterly basis. The total sick leave prevalence is roughly 2.7% in the central civil service and 3.2% in the private sector over the period. There seems to be a slightly rising trend in both sectors. It also depicts a substantial change between the year 2008 and the year 2009 in the private sector. In that same period, two disrupting events that might explain this change occurred. First, in July 2008, National Inter-professional Agreements (ANI) increased the generosity of the sickness benefit system in the private sector. This increased absenteeism in the private sector according to

---

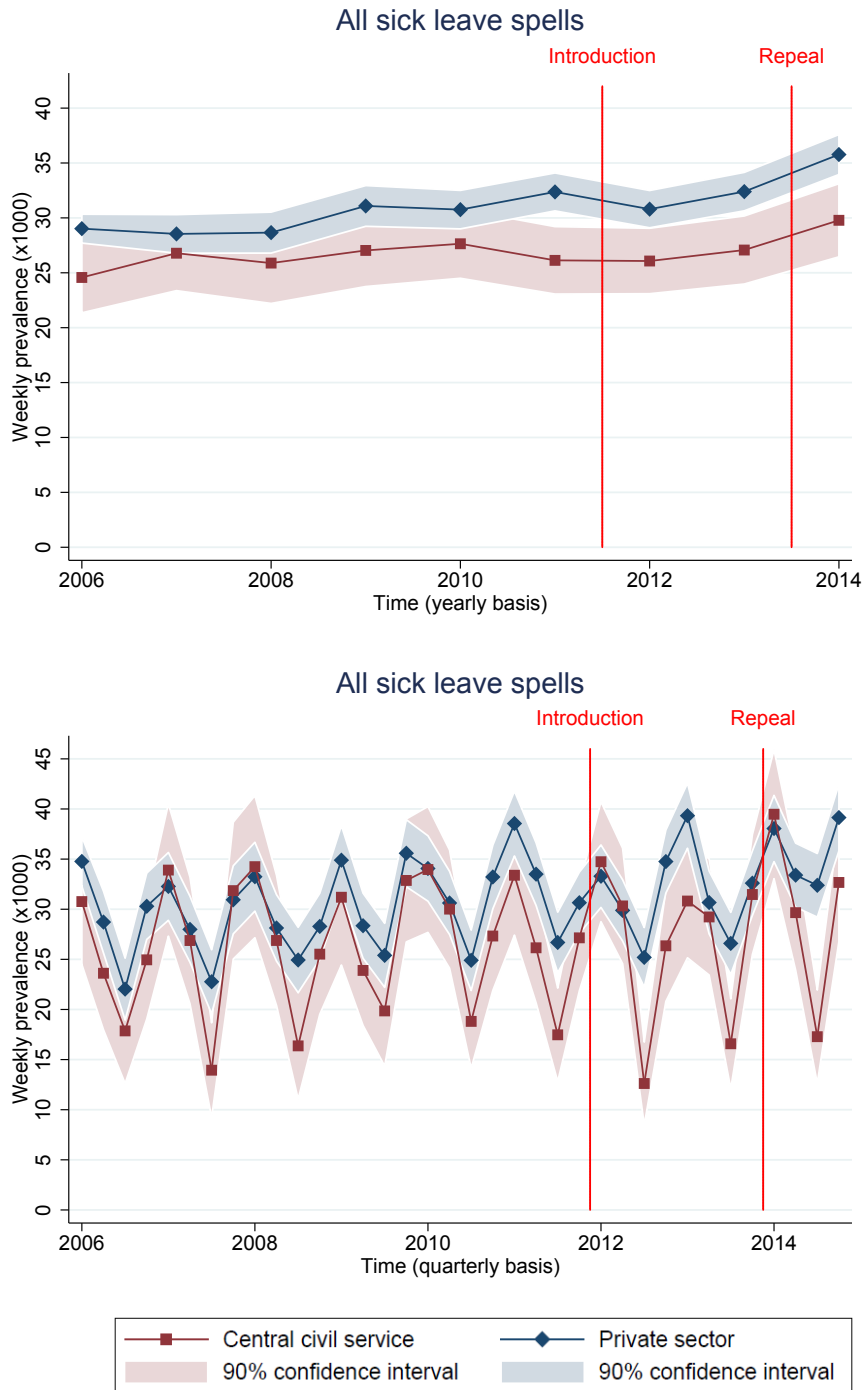
<sup>13</sup>Information on the precise part of the French civil service to which the employee may belong is available since 2006.

Ben Halima, Elbaz and Koubi (2016). Second, the 2008 crisis erupted and it may have affected the central civil service and the private sector differently between 2008 and 2009. As a consequence, we restrict the econometric analysis to years 2010-2014.

The two sectors seem to have evolved in a very similar way over the period regarding the prevalence of all sick leave spells, except between 2008 and 2009 for the two reasons mentioned above. This is the case until 2011, before the introduction of the policy, but also in 2012-2013, during the time of implementation of the policy, and in 2014, after the removal of the policy. At this stage, we have no evidence of an effect of the one-day waiting period.

**FIGURE 2.**

Prevalence of all sick leave by sector, at a yearly (top) and quarterly (bottom) basis.



Source: French Labour Force Survey 2006-2014.

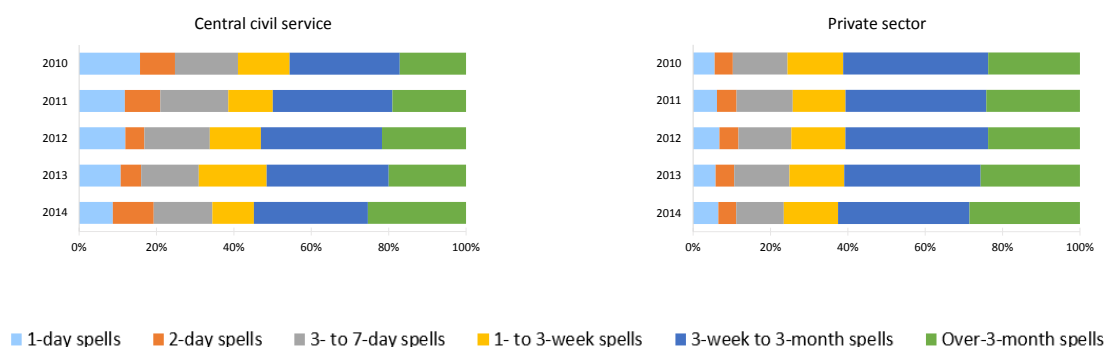
Lecture note: During the year 2010, the average weekly prevalence of all sick leave spells is 27.6‰ in the central civil service and 30.8‰ in the private sector. 90% confidence interval are computed using clustered-robust standard errors.



We go beyond the total prevalence by breaking sick leave spells down into the 6 previously described duration categories. We consider the distribution of observations of sick leave spells between these 6 categories for each year between 2010 and 2014 (Figure 3). Contrary to the total prevalence, a clear change appears at first sight. During the 2 years of the implementation of the policy (2012 and 2013), we observe a clear shift to the left of the spells distribution, in the treated group only. Looking more precisely, the category which decreases is the 2-day spells, while the category which increases is the 1- to 3-week spells. We thus continue by looking specifically at these two categories of sick leave spells,<sup>14</sup> to check whether this distribution change also comes with a level change for each of these two categories.

**FIGURE 3.**

Duration category distribution of ongoing sick leave spells by sector, years 2010-2014.



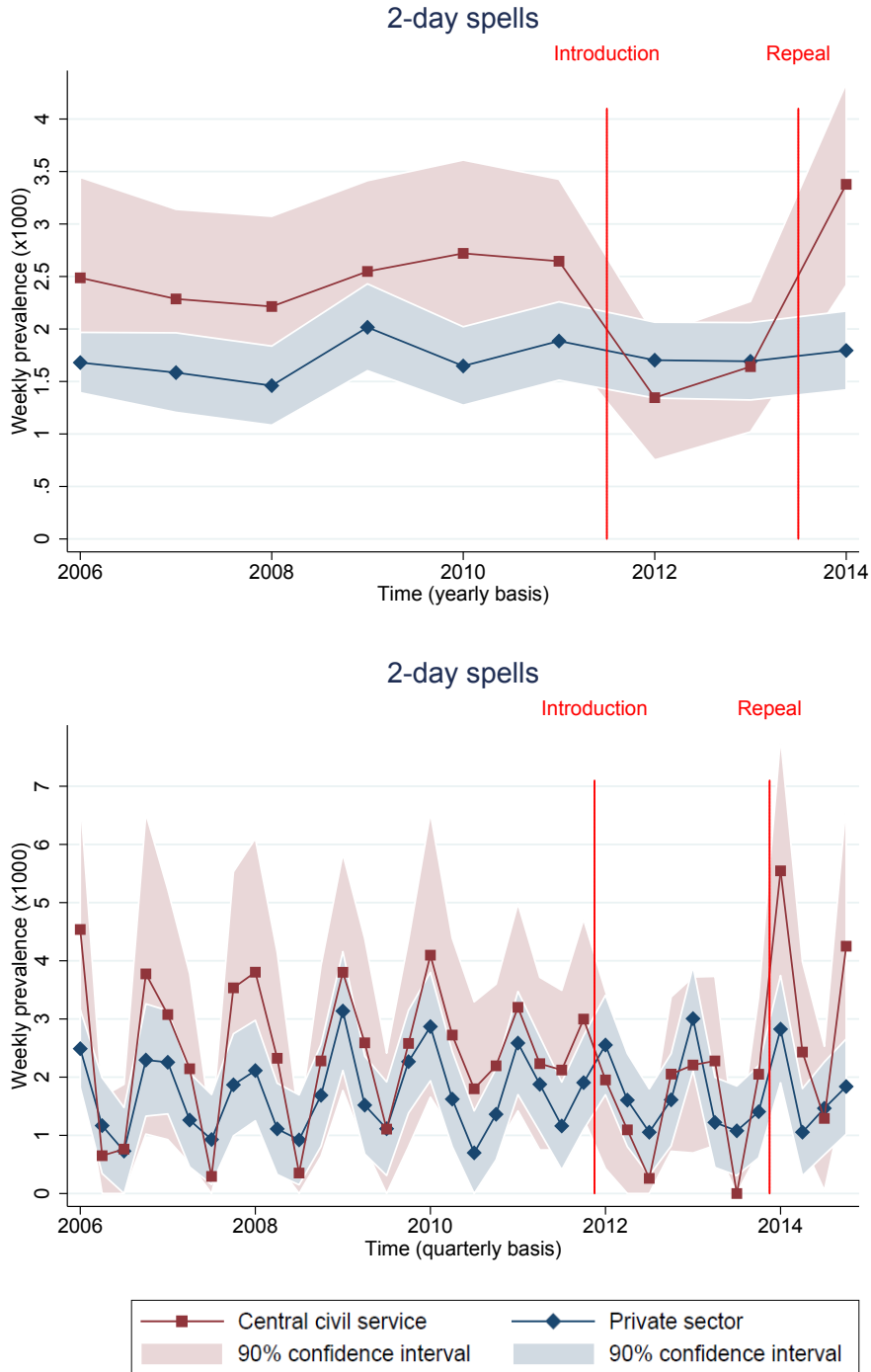
Source: French Labour Force Survey 2010-2014.

We first examine the prevalence of 2-day spells over time (Figure 4). We observe both a strong decrease at the time of introduction of the policy and a strong increase back to a pre-policy level at the time of repeal. The one-day waiting period seems to have strongly decreased the prevalence of 2-day spells, with an immediate effect at both the introduction and repeal.

<sup>14</sup>For the other categories, see Appendix B.

**FIGURE 4.**

Prevalence of 2-day spells by sector, at a yearly (top) and quarterly (bottom) basis.

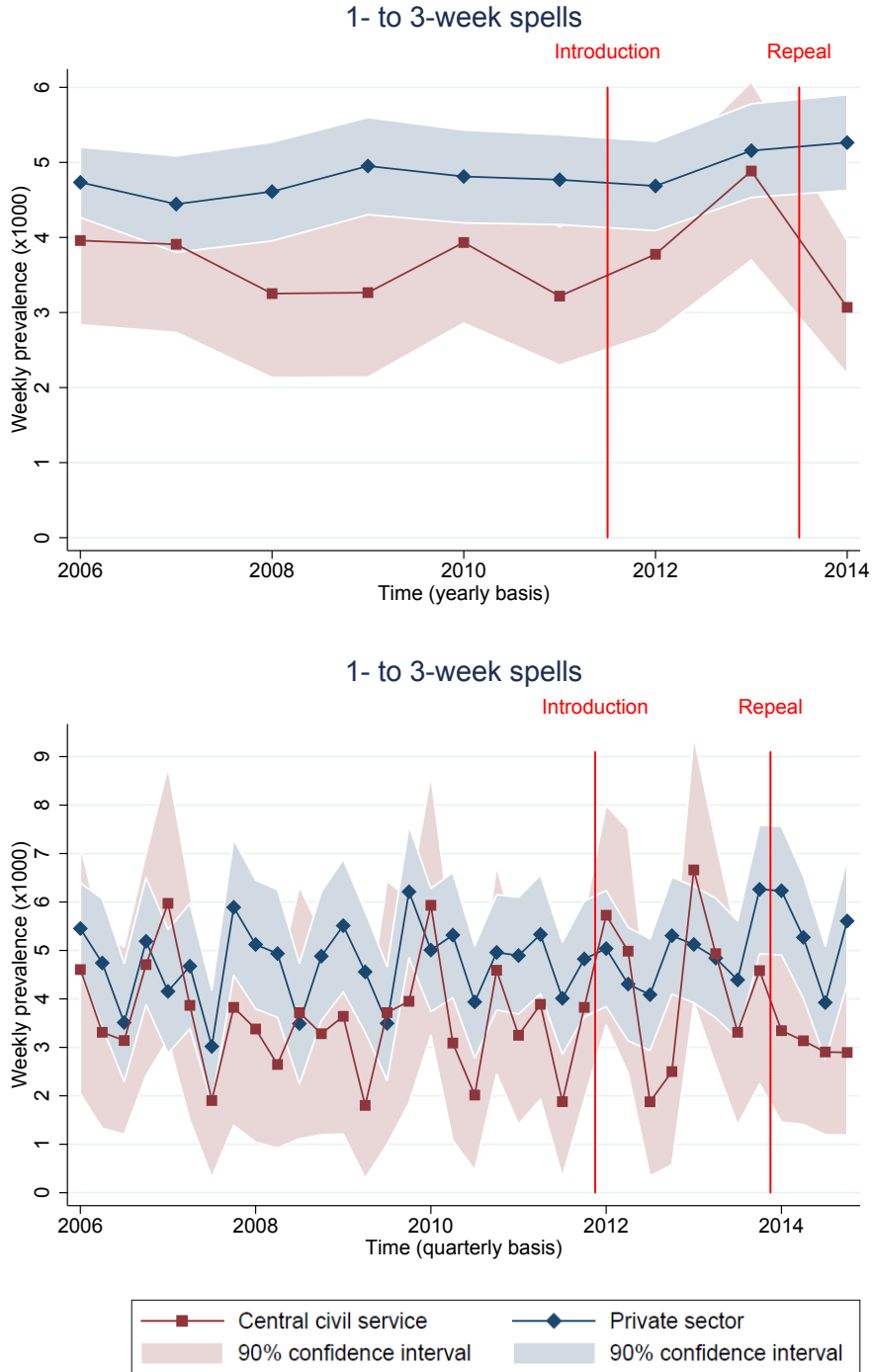


Source: French Labour Force Survey 2006-2014. Lecture note: During the year 2010, the average weekly prevalence of 2-day sick leave spells is 2.7‰ in the central civil service and 1.6‰ in the private sector. 90% confidence interval are computed using clustered-robust standard errors.

Second, we examine the prevalence of 1- to 3-week spells over time (Figure 5). We observe both an increase between the years 2012 and 2013, which is one year after the introduction of the policy, and a decrease back to a pre-policy level at the time of removal. The one-day waiting period seems to have increased the prevalence of 1- to 3-week spells, with a delayed effect at the introduction and an immediate effect at the repeal.

**FIGURE 5.**

Prevalence of 1- to 3-week spells by sector, at a yearly (left) and quarterly (right) basis.



Source: French Labour Force Survey 2006-2014.

Lecture note: During the year 2010, the average weekly prevalence of 1- to 3- week sick leave spells is 3.9‰ in the central civil service and 4.8‰ in the private sector. 90% confidence interval are computed using clustered-robust standard errors.

The global picture emerging from these descriptive statistics is hence a decreased prevalence of some short-term spells, an increased prevalence of some long-term spells, and an unchanged total prevalence. When comparing the means (see Table 1 below), the same picture emerges using temporal differences as well as difference-in-differences (see Section 3).

**TABLE 1.**  
Difference of prevalence by sector and by period:

<b>All sick leave:</b>			
Sector	Years 2010-2011-2014 [ $t_1$ ]	Years 2012-2013 [ $t_2$ ]	difference [ $t_2 - t_1$ ]
Private sector [ $s_1$ ]	32.980*** (0.432)	31.577*** (0.509)	-1.403** (0.644)
Public sector [ $s_2$ ]	27.909*** (1.066)	26.575*** (1.210)	-1.334 (1.552)
difference [ $s_2 - s_1$ ]	-5.072*** (1.150)	-5.002*** (1.313)	0.069 (1.680)

<b>2-day spells:</b>			
Sector	Years 2010-2011-2014 [ $t_1$ ]	Years 2012-2013 [ $t_2$ ]	difference [ $t_2 - t_1$ ]
Private sector [ $s_1$ ]	1.777*** 0.084	1.698*** 0.097	-0.079 0.127
Public sector [ $s_2$ ]	2.927*** 0.294	1.494*** 0.230	-1.434*** 0.373
difference [ $s_2 - s_1$ ]	1.150*** 0.306	-0.204 0.249	-1.355*** 0.394

<b>1- to 3-week spells:</b>			
Sector	Years 2010-2011-2014 [ $t_1$ ]	Years 2012-2013 [ $t_2$ ]	difference [ $t_2 - t_1$ ]
Private sector [ $s_1$ ]	4.950*** 0.143	4.916*** 0.169	-0.034 0.219
Public sector [ $s_2$ ]	3.398*** 0.294	4.331*** 0.436	0.932* 0.525
difference [ $s_2 - s_1$ ]	-1.551*** 0.327	-0.585 0.468	0.966* 0.569

*Source:* French Labour Force Survey 2010-2014.

*Lecture note:* Average prevalence of sick leave,  $\times 1000$ .

Cluster-robust standard errors in parentheses.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Simple statistics regarding health-related absences and sociodemographic characteristics for each sector over 2010-2014<sup>15</sup> are also presented in Table 2. The weekly prevalence

<sup>15</sup>For the analysis, we consider only the years 2010 to 2014 as explained above.

## 1.2. Institutional framework, data, and descriptive statistics

of all health-related absences is lower in the central civil service than in the private sector. The prevalence of very short-term absences (1- and 2-day spells) is higher in the central civil service, while the prevalence of all the other longer categories is lower. Most observable variables related to sociodemographic characteristics and working conditions are close. Employees of the central civil service are slightly more likely to be women, slightly older, slightly more likely to be in a couple and with a child under 6 years old. They are better paid and work slightly fewer days per week. The two main differences concern education and teachers. Concerning education, employees of the central civil service are much more likely to hold a graduate degree. The reason is that joining the central civil service in most cases requires applicants to pass a competitive exam, and enrolment requires them to hold various levels of degrees. Concerning teachers, most of them belong to the central civil service in France.

**TABLE 2.**

Health-related absences and observables characteristics by sector, years 2010-2014

Sector	Central civil service	Private Sector
<i>Weekly prevalences of sick leave, ×1000 (dependent variables)</i>		
All spells	27.4 ‰	32.4 ‰
1-day spells	3.48 ‰	2.29 ‰
2-day spells	2.37 ‰	1.75 ‰
3- to 7-day spells	4.56 ‰	4.91 ‰
1- to 3-week spells	3.76 ‰	4.94 ‰
3-week to 3-month spells	8.39 ‰	11.8 ‰
Over-3-month spells	4.83 ‰	6.75 ‰
<i>Sociodemographic characteristics and working conditions (covariates)</i>		
Women	54.4 %	45.4 %
Age (in years)	43.5	41.6
Highly educated (graduate degree level)	52.0 %	16.8 %
Being in a couple	74.5 %	73.7 %
Having a child under 6	21.5 %	20.1 %
Monthly wage	2302 €	1935 €
Teachers	45.7 %	0.53 %
Working days per week	4.70	4.88
Observations	70,779	518,176

*Source:* French Labour Force Survey 2010-2014.

From a cross section perspective, over the period we have 589,000 observations and among them 19,000 for which the respondent is on sick leave. From a panel perspective, we have 146,000 individuals. This implies that we have on average 4.03 observations per individual. The fact that this number is well below 6 is partially due to the definition

of the sample. First, we focus on the years 2010 to 2014, which leads to left- and right-censoring. Second, we impose a few restrictions on covariates, as explained in subsection 2.2. Considering only individuals whose first interview is at the earliest in 2010 Q1 and at the latest in 2013 Q3 (so that they can potentially be observed 6 times between 2010 Q1 and 2014 Q4), the average number of observations per individual rises to 5.2. This suggests that once an individual begins to participate in the survey, she answers most of the 6 interviews.

Among individuals of the sample, 132,000 are never on sick leave, 11,000 are only on sick leave once and 3,000 are on sick leave at least twice. When the individual is on sick leave at least twice, this most of the time implies at least one long-term spell (if the long-term spell lasts more than 3 months, it may be observed several times, as interviews are conducted quarterly). Only 145 individuals are at least twice on short-term sick leave.

Although some of the descriptive statistics are preliminary evidence of effects of the policy, we quantitatively identify the causal effects on each category of spells by carrying out an econometric analysis on our panel data. This serves to take into account time-variant observed and time-invariant unobserved individual heterogeneity regarding sick leave. In particular, the ability to control for the unobserved individual risk level allows us to conclude that the change in prevalence is due to a state dependence and not to an unobserved heterogeneity.

### **3 Empirical strategy**

To assess the effect of the presence of the one-day waiting period, we adopt a difference-in-differences strategy. The central civil service is the treated group and the private sector is the control group. Such public-private comparisons are common in this literature: there are for instance used by D'Amuri (2017), Pettersson-Lidbom and Thoursie (2013), and to some extent by Ziebarth and Karlsson (2010). The descriptive statistics above also showed that trends on the prevalence of sick leave in each sector were similar between the two groups before the introduction of the policy in January 2012.

Our main specification is the following:

$$y_{i,t} = \alpha.T_{i,t} + \beta.x_{i,t} + \mu_i + \nu_t + \epsilon_{i,t} \quad (1)$$

The dependent variable  $y_{i,t}$  is the dummy of employee  $i$  being absent from work for health-related reasons during her reference week of quarter  $t$ . It is a prevalence. Regressions are run for the dummy of all spells, but also for dummies of each of the 6 duration category spells, as described in Section 2.2.

The treatment dummy  $T_{i,t}$  is the presence dummy of the one-day waiting period for sick leave in the central civil service. The group dummy  $C_{i,t}$  is the dummy of employee  $i$  belonging to the central civil service at time  $t$ .  $T_{i,t}$  stands for the group dummy which is interacted with years 2012 and 2013:

$$T_{i,t} = C_{i,t} \times 1[2012 \text{ Q1} \leq t \leq 2013 \text{ Q4}] \quad (2)$$

$\alpha$  is the coefficient of interest. It captures the causal effect of the treatment, which is the presence of the one-day waiting period for sick leave.

$x_{i,t}$  stands for the group dummy  $C_{i,t}$ , as well as the socio-demographic and working conditions controls that may explain health-related absences and that are available in our data set. We include gender and age (through a spline function of age interacted with gender), a triple interaction of being in a couple, having a child under 6 and gender, the educational level and diploma, the professional category, the sector of activity, the type of contract, the number of working days during a usual week, the categorized paid vacation time per year, the housing occupation status and an interaction of the calendar quarter of the year with the sector.<sup>16</sup>

$\mu_i$  is an individual (employee) fixed effect. It controls for unobserved time-invariant individual-specific heterogeneity. Such fixed effects enable us to assess the effect of the policy using only the within variations. The inclusion of fixed effects increases the plau-

---

<sup>16</sup>Wage is not in the covariates, as we keep a unique value per employee, which is thus time-invariant.



sibility of estimation consistency in case non-response is affecting the common-trend assumption, as shown by Lechner, Rodriguez-Planas and Kranz (2016). This is almost never done in studies that use survey data and that are related to absenteeism. For instance, neither Ziebarth and Karlsson (2010, 2014), nor Goerke and Pannenberg (2015), nor D’Amuri (2017) use them. Puhani and Sonderhof (2010) use them in robustness tests that most often lose significance compared to their preferred specification. The impact of the inclusion of fixed effects is studied in subsection 5.3.

$\nu_t$  is the time effect of the quarter, from 2010 Quarter 1 to 2014 Quarter 4.<sup>17</sup>  $\epsilon_{i,t}$  is a heteroskedastic error term, whose variance is also clustered at the employee level.<sup>18</sup>

## 4 Results

### 4.1 Treatment effects for spells of different durations

Table 3 presents the results for the main specification. The regressions yield results that confirm what was suggested by the descriptive statistics. Regarding the prevalence of all sick leave spells, no effect is found.

**TABLE 3.**  
Treatment effects on the prevalence of spells ( $\times 1000$ )

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
T	-0.0240 (0.879)	-1.845*** (0.661)	0.469 (0.970)	1.840** (0.865)	0.806 (1.518)	1.087 (1.303)	2.333 (2.466)
Observations	588,955	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00161	0.00188	0.00196	0.00113	0.00274	0.00408	0.00399

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

<sup>17</sup>The time dummy  $1[2012\text{ Q1} \leq t \leq 2013\text{ Q4}]$  is not included, as it would be collinear with the quarterly effects  $\nu_t$ .

<sup>18</sup>This allows us to address any downward bias in the standard errors due to serial correlation, as highlighted by Bertrand, Duflo and Mullainathan (2004).

The coefficient regarding the prevalence of 2-day spell is negative and highly significant, but this is not the case for the two other categories of spells inferior to one week. The coefficient for 2-day spells has to be compared to the mean value of 2-day spells in the central civil service when the policy is not in place, which is 2.93‰. This implies a decrease of 63% of 2-day spells due to the one-day waiting period. For 1-day spells, the coefficient is not significant. The absence of a negative effect might come from their shortness, which makes it particularly easy to report them to the employer as leave for other reason than sick leave, like days off,<sup>19</sup> or not to report them. Due to this layer of hidden action regarding the reporting of sick leave, very short health-related absences can remain unaffected by the one-day waiting period. We would expect a negative coefficient if we were considering 1-day sick leave (as declared to the employer) rather than 1-day health-related absence. For 3- to 7-day spells, the coefficient is also not significant. Due to the one-day waiting period, employees may be reluctant to begin sick leave. However, when the sick leave is taken, its duration may increase costlessly. For example, sick leave which would last 2 days lasts 3 days. Similarly, sick leave which would last 3 days lasts 4 days, and so on. The impact on the prevalence of intermediate duration is therefore unclear. This may explain why we do not get a clear-cut effect on these spells while we get one on 2-day spells.

Regarding the prevalence of long-term spells, the three coefficients are positive. It is significant for 1- to 3-week spells, as suggested by the descriptive statistics. The prevalence of 1- to 3-week spells is 3.40‰ in the central civil service when the policy is not in place. It implies an increase of 54% due to the one-day waiting period. It is non-significant for the two categories of spells over 3 weeks, which may concern mostly serious diseases. Overall, since the prevalence of long-term spells is higher than the prevalence of short-term ones, the prevalence of all sick leave spells is driven more by the long-term spells, which explains why the coefficient for all sick leave spells is positive.

To gather further insights, we separately interact the treatment dummy with years

---

<sup>19</sup>In some administrations, such as customs, there even exists leave for very short indispositions. In the particular case of customs, the French highest Court of Administrative Justice (Conseil d'État) stated that the one-day waiting period did not apply to this leave for very short indispositions (Conseil d'État, 2013).

**TABLE 4.**Treatment effects on the prevalence of spells ( $\times 1000$ ), by year of implementation of the policy

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
T $\times$ Year 2012	0.268 (1.079)	-2.029*** (0.757)	0.924 (1.152)	0.960 (1.067)	0.823 (1.620)	1.739	2.686
T $\times$ Year 2013	-0.325 (0.983)	-1.655** (0.773)	-0.000113 (1.153)	2.746** (1.076)	0.788 (1.900)	0.415 (1.596)	1.970 (2.989)
Observations	588,955	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00161	0.00188	0.00196	0.00113	0.00274	0.00408	0.00399

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

2012 and 2013. Results are presented in Table 4. This more detailed table yields overall the same results as the previous one. The two coefficients for 2-day spells are still very significant. The effects are also of the same magnitude each year. The main difference with the previous table lies in the coefficient for 1- to 3-week spells, which is lower in 2012 than in 2013. It is not significant in 2012, whereas it remains significant in 2013. This latter point might indicate that the increase in the prevalence of long sick leave did not take place immediately after the implementation of the policy, but required some time to reach its full effect, as seen in the descriptive statistics.

We also present in Appendix D the interactions of the treatment dummy with each quarter of 2012 and 2013 (see Table D.I). The coefficients for 2-day spells are always negative and are significant each Quarter 1 and Quarter 3. The coefficient for 1-day spells and for 3 to 7 days spells are sometimes positive, sometimes negative and are never significant. The coefficients for 1- to 3-week spells are all positive save one, and two of them are significantly positive in the first two quarters of 2013, during the second year of implementation of the policy. The coefficient for longer absences are sometimes positive, sometimes negative and all save one are not significant. Overall, these results corroborate what was found in Table 4.

The global picture put forward by the descriptive statistics is therefore confirmed by the econometric analysis. The presence of the one-day waiting period induces a decreased

prevalence of short-term absences, an increased prevalence of long-term absences, and an unchanged total prevalence. The underlying mechanism suggested by these results is a decreased incidence of sick leave and an increased duration of spells. This is consistent with what was found by Pettersson-Lidbom and Thoursie (2013) in a similar context in Sweden where a one-day waiting period was repealed. Their "*major point [...] is that the reform made individuals start new spells to a larger extent but that ongoing spells became shorter*".

The explanation of the decreased incidence is straightforward: the cost of starting a spell has a deterring effect on doing so. Explanations of the increased duration of spells are subtler. Intuitively, we have in mind three theoretical mechanisms. First, a rational forward-looking explanation. Starting a new sick leave spell implies paying only a fixed cost, for sick leave up to 3 months. Once a spell is started and the first fixed cost is paid, a forward-looking employee prefers to stay longer on sick leave so that the probability of getting sick again decreases, in order to avoid paying this fixed cost a second time. This results in an increased duration of spells. This explanation is put forward by Johansson and Palme (2005), Paola, Scoppa and Pupo (2014) and Eliason, Johansson and Nilsson (2019). Second, a health capital explanation. The one-day waiting period deters the employee from starting a sick leave. Consequently, the policy would induce a degradation of health capital. After a certain time period, sick employees are forced to stop. Their sick leave spells are longer, due to a poorer state of health. Such an adverse consequence of presenteeism would explain<sup>20</sup> why the increase in 1- to 3-week spells is observed only with a delay, contrary to the decrease in 2-day spells which is observed instantaneously. As for the first explanation, this second one is dynamic. But contrary to the first explanation, the increased duration is suffered, and not chosen. Third, there is a behavioural and static explanation. The one-day waiting period may be seen as a deductible: the employee pays the cost of the first day, but nothing else until 3 months in our setting. This deductible may be seen as unfair. Consequently, the employee may be tempted to compensate for this perceived unfairness by increasing the duration of her

---

<sup>20</sup>Note that this explanation is not in line with the results of the articles that have studied a 1997 German reform. If health capital mattered at that time in Germany too, the two opposite changes in the distribution for short and long spells would arguably also be present in this simpler reform with a uniform decrease in the replacement rate. Yet, no health degradation was observed by the studies that carefully examined it (Ziebarth and Karlsson, 2014; Puhani and Sonderhof, 2010).

sick leave spell, to enjoy more leisure time.<sup>21</sup> The fact that the measure was continuously and unanimously criticized as unfair by labour unions provides credible grounds for this explanation. This suggests that this feeling of unfairness was widespread and might have been sustained over the two years of implementation of the measure. The effects we found are also in line with a small body of empirical literature focusing on deductibles in car insurance (Dionne and Gagné, 2001; Miyazaki, 2009; von Bieberstein and Schiller, 2017), which finds that a higher deductible may lead to a higher reported cost of car crashes.<sup>22</sup>

## 4.2 Heterogeneous effects

In the following, we explore the possible heterogeneous effects by covariate. We interact the treatment dummy with various covariates related to socio-demographic characteristics and working conditions.<sup>23</sup> Results are presented in Table 5. Each column summarizes a unique regression. A coefficient can be interpreted as the effect of the associated covariate on the intensity of the reaction, with the effect of other covariates on the intensity being held constant, even if there is a correlation between these covariates.

Three covariates significantly change the intensity of the reaction: two related to sociodemographic characteristics (gender and age) and one related to working conditions (the number of working days during a usual week<sup>24</sup>). These three covariates change the intensity of the decrease in 2-day spells: being a woman, being younger or usually working

---

<sup>21</sup>Voss, Floderus and Diderichsen (2001) refer to a Danish study which has similar results to ours and puts forward this explanation: "In Denmark, the introduction of a [qualifying day] in 1983 was followed by a clear decrease in short-term sick-leave events (1-3 days) and to some extent an increase in longer sick-leave events (>4 days). One explanation from the authors was that some people might compensate with an extra day of sick-leave if the [qualifying day] was experienced as unreasonable."

<sup>22</sup>Dionne and Gagné (2001) do not mention how the cost of car accidents per individual is affected by a higher deductible, but only that the cost per car accident increases. In our case, we have more detailed results. Not only are our findings consistent with a longer spell duration (whose equivalent in car insurance is an increased cost per car crash), but also we focus on the total prevalence (whose equivalent in car insurance is the average cost per individual) and we do not find a significant decrease. Hence, the classic deterrent effect of the deductible on the incidence of sick leave (whose equivalent in car insurance is the probability of claims) would be fully offset by the increased duration of sick leave spells.

<sup>23</sup>Most covariates are significantly correlated with absenteeism levels over the studied period. See Table C.I in Appendix C for pooled OLS regressions over the 2010-2014 period.

<sup>24</sup>In the survey, both the number of working days during a usual week and during the reference week are available. We use the former to decrease concerns of endogeneity with the treatment. More generally, the need to use non-endogenous covariates led us to consider working conditions during a usual week and not during the reference week. This drastically decreases the number of suitable covariates.

**TABLE 5.**  
Heterogeneous treatment effects on the prevalence of spells ( $\times 1000$ )

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
T	8.601 (5.667)	-15.55*** (4.961)	5.850 (6.755)	-0.224 (4.924)	22.00* (11.57)	3.931 (8.111)	24.61 (16.41)
T $\times$ Women	0.476 (1.758)	-2.394** (1.084)	-1.914 (1.771)	2.727* (1.636)	-4.275 (2.952)	-1.575 (2.199)	-6.956 (4.538)
T $\times$ Age	-0.173** (0.0823)	0.141** (0.0690)	-0.151 (0.100)	0.0690 (0.0917)	-0.226 (0.167)	-0.0699 (0.160)	-0.410 (0.276)
T $\times$ Working days during a usual week	0.115 (1.011)	1.696** (0.781)	0.220 (1.196)	-0.208 (0.833)	-2.783 (1.863)	-0.312 (1.324)	-1.271 (2.789)
T $\times$ Wage	-0.425 (0.549)	0.518 (0.356)	0.150 (1.054)	-0.452 (0.885)	0.938 (1.832)	1.279 (1.181)	2.009 (2.550)
T $\times$ Teacher	-1.802 (1.800)	-0.749 (1.160)	1.911 (1.747)	-0.892 (1.639)	4.125 (2.843)	-0.913 (2.320)	1.681 (4.586)
Observations	588,955	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00164	0.00194	0.00197	0.00113	0.00276	0.00408	0.00401

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Note: Wage is expressed in thousand euros per month.

fewer days per week independently implies a higher decrease. In addition, being a woman also increase the intensity of the increase in 1- to 3-week spells: women who hence tend to react more, regarding both short-term and long-term spells. Moreover, old employees would tend to decrease 1-day spells instead of 2-day spells, compared to young employees, as the two corresponding coefficients have opposite signs but similar absolute value.

Regarding gender differences, the two reactions are higher for women, though the change in total prevalence is eventually not significant, and its sign points towards a higher reaction for men. In the literature, results on gender differences regarding behavioural responses to monetary incentives for sickness absence are far from being unanimous. Some authors find that men react more strongly than women (Johansson and Palme, 2005; Ziebarth and Karlsson, 2014) whereas others find the opposite (Paola, Scoppa and Pupo, 2014; Pettersson-Lidbom and Thoursie, 2013) or find no difference (Puhani and Sonderhof, 2010).

Several factors might explain why responses to the one-day waiting period policy along with age, gender and number of working days during a usual week are heterogeneous.

For age, young employees may be more ready than old employees to report 1-day health-related absences to their employers as leave for other reasons than sick leave, avoiding consequently to take 2-day health-related absences (less easy to report as leave for other reasons than sick leave). Young employees would hence make more use of their reporting choice. For gender, we find that the policy affected more the prevalence of short-term spells and of longer spells for women. In other contexts, different conclusions were drawn for gender, which calls for caution with the external validity of these results. For the number of working days during a usual week, the explanation may be that a shorter working week increases the degrees of freedom of the employee regarding her weekly schedule. It is easier for her to shift the day dedicated to illness recovery to a non-working day.

Wage is not a determinant of the intensity of the reaction. Although the one-day waiting period is a monetary incentive, the penalty is proportional to the wage. The relative loss is thus identical whatever the wage. This may explain why employees with different wages do not react differently, other observable characteristics being held constant.<sup>25</sup>

### **4.3 Switch on and switch off**

A notable feature of the policy under study is its short period of enforcement, since it was repealed exactly two years after its implementation. As noted by Puhani and Sonderhof (2010), such a feature is of particular interest in a difference-in-differences approach. Indeed, this approach relies on the common trend assumption. Without treatment, the two groups are assumed to evolve similarly. With a single switch, it can be difficult to test whether any estimated effect could come from a violation of this assumption. For instance, even if the studied policy had no real effect, a significant estimated effect might come from another simultaneous event going in the same direction, and of which the researchers would be unaware. When the studied policy implies both a switch on and a switch off, we can be more confident in the fact that we have effectively assessed the causal impact of the policy, if we find that both estimated effects have opposite signs and are of similar magnitude. In the latter example, it is indeed more unlikely that for

---

<sup>25</sup>When the treatment dummy interacted with the wage is the sole interacted term, a lower wage implies a significantly stronger decrease in 2-day spells. This effect is however due to the correlation of the wage with other covariates, such as age.

both the switch on and the switch off, there would be two simultaneous events going in opposite directions, and of which the researchers would be unaware.

**TABLE 6.**

Treatment effects of the switch on and of the switch off (prevalence  $\times$  1000)

	2-day spells		1- to 3-week spells	
	Switch on (2012 <i>versus</i> 2011)	Switch off (2014 <i>versus</i> 2013)	Switch on (2012 <i>versus</i> 2011)	Switch off (2014 <i>versus</i> 2013)
T	-1.441 (0.994)	2.365** (0.936)	1.282 (1.335)	-1.878 (1.346)
Observations	244,322	226,795	244,322	226,795
Individuals	74,368	72,934	74,368	72,934
$R^2$	0.00415	0.00343	0.00293	0.00219

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Using two successive years, effects estimated separately on the switch on and the switch off are respectively presented in Table 6, for 2-day and 1- to 3-week spells, the two most reactive categories. At first glance, all coefficients have the expected sign. For 2-day spells, the coefficients are negative between 2011 and 2012 and positive between 2013 and 2014. Only the switch off is significant due to its stronger magnitude,<sup>26</sup> which might come from the transitory peak in 2-day spells of the central civil service in the quarter following the abolishment of the measure (see Figure 4). For 1- to 3-week spells, though the coefficients are not significant, they are of expected sign: positive between 2011 and 2012, negative between 2013 and 2014. Moreover, they are of similar magnitude.

## 5 Robustness tests

### 5.1 Placebo test

Similarly to what was conducted regarding the switch on and the switch off, an additional robustness check is to test whether the time pattern is similar in the two sectors during

<sup>26</sup>This stronger magnitude hints that the one-day waiting period might have effects in the period following its abolishment, perhaps by decreasing the prosocial motivation of civil servants (Lanfranchi and Lemoyne, 2020). The *difference* in magnitude is however not significant in our analysis.



the pre-reform period (between 2010 and 2011). Such tests are often used to dismiss the existence of diverging trends that may bias the results. Under the common trend assumption, we expect to find no significant results in these tests.

**TABLE 7.**  
Treatment effects in placebo tests (prevalence  $\times$  1000)

	2-day spells	1- to 3-week spells
Placebo T (2011 <i>versus</i> 2010)	-0.568 (1.114)	-0.495 (1.491)
Observations	240,933	240,933
Individuals	73,554	73,554
$R^2$	0.00385	0.00309

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Results are presented in Table 7. Regarding 2-day and 1- to 3-week spells, coefficients are either positive or negative and they are never significant. They are also much lower than those found in Table 6.

## 5.2 Alternative control group

The difference-in-differences strategy relies on the assumption that the control group is a good counterfactual for the treated group. Table 2 has shown that most covariate means are of the same order between the two groups, but that a few observable characteristics, such as the educational level, differ. To get a more similar control group in terms of socioeconomic characteristics and working conditions without losing too many observations, we consider the private service sector as an alternative control group. Results are presented in Table 8. They show a similar magnitude. Standard errors are higher, which is consistent with the use of a lower number of observations.

**TABLE 8.**Treatment effects with alternative control group (prevalence  $\times$  1000)

Control group	2-day spells		1- to 3-week spells	
	Private service sector	Whole private sector (Baseline)	Private service sector	Whole private sector (Baseline)
T	-1.816*** (0.681)	-1.845*** (0.661)	2.048** (0.896)	1.840** (0.865)
Observations	427,151	588,955	427,151	588,955
Individuals	107,038	146,035	107,038	146,035
$R^2$	0.00202	0.00187	0.00162	0.00112

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

### 5.3 Covariates and fixed effects

As robustness tests, we consider alternative specifications for our two main results, the effect of the policy for 2-day and 1- to 3-week spells. The results of those specifications are presented in Table 9. We remove on the one hand the individual fixed effects, and on the other hand the covariates.

**TABLE 9.**Treatment effects with different econometric specifications (prevalence  $\times$  1000)

Specification	2-day spells			1- to 3-week spells		
	Pooled OLS with covariates	FE without covariates	FE with covariates (Baseline)	Pooled OLS with covariates	FE without covariates	FE with covariates (Baseline)
T	-1.381*** (0.395)	-1.888*** (0.660)	-1.845*** (0.661)	1.106* (0.568)	1.855** (0.865)	1.840** (0.865)
Observations	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00165	0.000362	0.00187	0.00223	0.000147	0.00112

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Removing the fixed effects decreases the magnitude of the two effects in absolute value. However, both remain significant. When OLS and FE estimations differ on an unbalanced panel, Lechner, Rodriguez-Planas and Kranz (2016) suggest that it "should be considered as evidence that non-response is not ignorable for the differences-in-differences estimation". In the survey, we saw in subsection 2.3, that once an individual begins to

respond to the survey, she responds to most of the six interviews. Even though answering the survey is mandatory, it might be suggested that the propensity to begin to respond to the survey is decreasing with health status. If the one-day waiting period has a negative impact on health status, affected employees would choose not to respond to the survey. This would explain why the increase in 1- to 3-week spells is lower when considering the OLS estimation instead of the FE one. If non-response is not ignorable, then the FE estimation is to be preferred to the OLS one, as only the former may be consistent. On the contrary, not including the covariate has almost no impact on the results. Time variant controls are not likely to bring much information in addition to the time invariant employee fixed effect, as we observe individuals during at most 6 quarters.

## 5.4 Weights

In all the previous regressions, weights were used. Though the use of weights is clearly needed to get proper descriptive statistics from a survey, there is a debate among statisticians about whether they should be used for regressions (see Solon, Haider and Wooldridge (2015)). One way to deal with this issue consists in comparing regressions with and without weights, in order to see the extent to which it matters. In our case, results are not very sensitive to the inclusion of the weights (see Table 10).

**TABLE 10.**  
Treatment effects without and with using the weights (prevalence  $\times$  1000)

Weighting Trimming	2-day spells			1- to 3-week spells		
	No No	Yes Yes (Baseline)	Yes No	No No	Yes Yes (Baseline)	Yes No
T	-2.097*** (0.682)	-1.845*** (0.661)	-1.707*** (0.645)	1.568* (0.840)	1.840** (0.865)	2.249** (1.006)
Observations	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00203	0.00187	0.00193	0.00100	0.00112	0.00114

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 5.5 Nonlinear specifications

Despite the advantages of our main linear specification,<sup>27</sup> a first concern is that it may be not fully adapted to situations in which a dummy dependent variable is close to zero, as is the case here. A more natural model is a logit model. Compared to the linear probability model, the two may have similar behaviours in the linear zone of the logistic function, when the dependent variable is close to 0.5, but not when the dependent variable is close to 0 or 1.

We therefore perform a robustness check with a logit model instead of a linear probability model. Using fixed effects in a logit model makes the estimation subject to the incidental parameter problem. When the number of periods is small, estimators are inconsistent. A solution lies in considering a conditional fixed effect logit model (Chamberlain, 1980).<sup>28</sup> The computation of average treatment effects is not straightforward, as fixed effects are not estimated. We can however interpret the sign of the coefficient of the interaction term as the sign of the treatment effect (see Puhani (2012)).

We present the results of the conditional fixed effect logit model in Table 11. We do not present the raw coefficients but their exponential, as the latter can be interpreted as relative changes. At first sight, we have smooth decreasing (from 1-day to 2-day spells) coefficients and then an increasing pattern. Numerically, relative changes are straightforward: 2-day spells decrease by 57%, while 1- to 3-week spells increase by 49%. These relative changes are close to those computed with the linear probability model (see Section 4) and they are significant. Conversely, 1-day spells decrease in this specification, but the associated coefficient remains non-significant.

A second concern is related to the relationship between the different probabilities we consider. We partition the possible events into different categories and we consider separately the probabilities of each event, without taking into account the fact that the sum

---

<sup>27</sup>We can cite the ability to directly interpret coefficients as treatment effects, the ability to interact the treatment dummy with covariates and to easily interpret the coefficients of the interacted terms, computational ease (no numerical problem of convergence as for likelihood maximization), the ability to easily cluster standard error to avoid downward biases of the standard errors due to serial correlation.

<sup>28</sup>Weights are not included as there are not available in the command we use - Section 5.4 however showed that in the linear case, results were not very sensitive to the inclusion of weights.

**TABLE 11.**  
Binary logit with fixed effects

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
T	0.725 (0.249) [0.468,1.124]	0.432** (0.304) [0.270,0.692]	1.003 (0.214) [0.686,1.466]	1.489* (0.220) [1.038,2.136]	1.121 (0.152) [0.827,1.519]	1.269 (0.229) [0.788,2.046]	1.017 (0.0915) [0.849,1.218]
Observations	6,484	5,035	13,385	12,948	28,149	11,283	63,750
Individuals	1,306	1,012	2,685	2,589	5,695	2,274	12,867
Pseudo $R^2$	0.0380	0.0464	0.0278	0.0110	0.0142	0.0613	0.0170

Odds ratios are displayed. Cluster-robust standard errors in parentheses.

There are estimated through 50 bootstrap replications. 90% confidence interval below.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Note: Conditional fixed effect multinomial logit model only uses employees who experience a change in the dependent variable, which explains why the number of observations is low compared to the linear probability model.

of these probabilities, including that of not reporting a sick leave, is equal to one. To investigate whether imposing this condition matters, we consider a multinomial conditional fixed effect logit model. Each modality is one of the 6 considered duration categories, and the reference modality is not having a sick leave. Numerical implementation is made possible by Pforr (2014). Results are presented in Table 12. We find results which are similar both to the binary conditional fixed effect logit model and to the linear probability model.

**TABLE 12.**  
Multinomial logit with fixed effects

	Odds ratio	Standard error	90% confidence interval
1-day spells	0.728	(0.208)	[0.456,1.165]
2-day spells	0.432***	(0.114)	[0.280,0.667]
3- to 7-day spells	1.012	(0.150)	[0.794,1.292]
1- to 3-week spells	1.524*	(0.343)	[1.052,2.208]
3-week to 3-month spells	1.162	(0.192)	[0.886,1.526]
Over-3-month spells	1.323	(0.371)	[0.834,2.099]
Observations		64,636	
Individuals		13,130	
Pseudo $R^2$		0.0307	

Odds ratios are displayed. Cluster-robust standard errors in parentheses.

There are estimated through 50 bootstrap replications. 90% confidence interval below.

Note: Conditional fixed effect logit models only uses employees who experience a change in the dependent variable, which explains why the number of observations is low compared to the linear probability model.

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 6 Conclusion

Our results provide further support for the thesis that the pattern of sick pay matters for health-related absences. They imply the existence of a hidden action of the agent which depends on the cover framework. We thus conclude on the presence of moral hazard. But it is only the duration distribution of sick leave that is impacted, not the total prevalence.

Whereas the usual theoretical prediction in a static framework is that there is a trade-off between coverage and incentive, we do not find that the one-day waiting period, which represents less generous coverage, leads to a decrease in the total prevalence of sick leave. If anything, it increases total prevalence. We find that this monetary incentive alters the duration distribution of sick leave spells. The policy leads to a significant decrease in the prevalence of 2-day spells and to a significant increase in the prevalence of 1- to 3-week spells. The corresponding changes are of a large magnitude. We find a decrease of more than half in 2-day spells and an increase of a half in 1- to 3-week spells. These two effects of the one-day waiting period go into opposite directions, which results in the stability of the total prevalence level. Whether the one-day waiting period may or may not have any impact on productivity is unclear, since the partition of the same level of absence between short-term absence and long-term absence could also matter to this respect.

In addition to these findings, we also document heterogeneous effects of the one-day waiting period along with age, gender, number of working days per week. Being a young employee, a woman or working fewer days per week implies a higher decrease in 2 days sick leave ; and being a woman implies a higher increase in 1- to 3 week sick leave. This might suggest that the sick pay pattern varies with certain sociodemographic characteristics or working conditions.

Our main findings are consistent with results previously found in few other papers based on quasi-natural experiments and focusing on the existence of a waiting period (Davezies and Toulemon, 2015; Pettersson-Lidbom and Thoursie, 2013; Voss, Floderus and Diderichsen, 2001) or on the implementation of similar schemes involving the replacement rate (Johansson and Palme, 2005; Paola, Scoppa and Pupo, 2014). We thus contribute to the assessment of the external validity of these results. It seems that such

opposite effects regarding the prevalence of short-term spells on the one hand and long-term spells on the other are observed when the change introduces a locally increasing marginal coverage which peaks at 100%. Conversely, these opposite effects are not observed when the coverage remains monotonically decreasing with the duration of sick leave (as in Ziebarth and Karlsson (2010, 2014); Puhani and Sonderhof (2010); Goerke and Pannenberg (2015)) or when the replacement rate is strictly below 100% after the initial period (as in Pertold (2019)).

---

## References

- Ben Halima, Mohamed Ali, Nathan Elbaz, and Malik Koubi.** 2016. “The effect of expanding the generosity of the statutory sick leave insurance: the case of a French reform.” *mimeo*.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan.** 2004. “How Much Should We Trust Differences-In-Differences Estimates?\*” *The Quarterly Journal of Economics*, 119(1): 249.
- Cazenave-Lacroutz, Alexandre, and Alexandre Godzinski.** 2017. “Effects of the one-day waiting period for sick leave on health-related absences in the French central civil service.” Institut National de la Statistique et des Etudes Economiques.
- Chamberlain, Gary.** 1980. “Analysis of Covariance with Qualitative Data.” *The Review of Economic Studies*, 47(1): 225–238.
- Chemin, Matthieu, and Etienne Wasmer.** 2009. “Regional Difference-in-Differences in France Using the German Annexation of Alsace-Moselle in 1870-1918.” *NBER Chapters*, 5(1): 285–305.
- Conseil d’État.** 2013. “Décision numéro 358896 rendue le 04/10/2013 par les 3ème et 8ème sous-sections réunies.” <https://www.legifrance.gouv.fr/affichJuriAdmin.do?idTexte=CETATEXT000028036207>.
- D’Amuri, Francesco.** 2017. “Monitoring and disincentives in containing paid sick leave.” *Labour Economics*, 49: 74–83.
- Davezies, Laurent, and Léa Toulemon.** 2015. “Does Moving to a System with a More Generous Public Health Insurance Increase Medical Care Consumption?” *Annals of Economics and Statistics*, 119-120: 179–205.
- DGAFP.** 2015. “Rapport annuel sur l’état de la fonction publique - édition 2015.” [https://www.fonction-publique.gouv.fr/files/files/statistiques/rapports\\_annuels/2015/RA\\_2015\\_definitif.pdf](https://www.fonction-publique.gouv.fr/files/files/statistiques/rapports_annuels/2015/RA_2015_definitif.pdf).



- Dionne, Georges, and Robert Gagné.** 2001. “Deductible contracts against fraudulent claims: evidence from automobile insurance.” *Review of Economics and Statistics*, 83(2): 290–301.
- Drees.** 2014. “La protection sociale en France et en Europe en 2012 - édition 2014.” <http://www.drees.sante.gouv.fr/IMG/pdf/er888.pdf>.
- Eliason, Marcus, Per Johansson, and Martin Nilsson.** 2019. “Forward-looking moral hazard in social insurance.” *Labour Economics*, 60: 84–98.
- Goerke, Laszlo, and Markus Pannenberg.** 2015. “Trade union membership and sickness absence: Evidence from a sick pay reform.” *Labour Economics*, 33: 13–25.
- Henrekson, Magnus, and Mats Persson.** 2004. “The effects on sick leave of changes in the sickness insurance system.” *Journal of Labor economics*, 22(1): 87–113.
- Ichino, Andrea, and Regina T Riphahn.** 2005. “The effect of employment protection on worker effort: Absenteeism during and after probation.” *Journal of the European Economic Association*, 3(1): 120–143.
- Johansson, Per, and Mårten Palme.** 2005. “Moral hazard and sickness insurance.” *Journal of Public Economics*, 89(9): 1879–1890.
- Lanfranchi, Joseph, and Priscilla Lemoyne.** 2020. “Motivation prosociale et présentéisme dans les secteurs public et privé.” *Revue économique*, 71(6): 1101–1132.
- Lechner, Michael, Nuria Rodriguez-Planas, and Daniel Fernández Kranz.** 2016. “Difference-in-difference estimation by FE and OLS when there is panel non-response.” *Journal of Applied Statistics*, 43(11): 2044–2052.
- Miyazaki, Anthony D.** 2009. “Perceived ethicality of insurance claim fraud: Do higher deductibles lead to lower ethical standards?” *Journal of Business Ethics*, 87(4): 589–598.
- Paola, Maria De, Vincenzo Scoppa, and Valeria Pupo.** 2014. “Absenteeism in the Italian Public Sector: The Effects of Changes in Sick Leave Policy.” *Journal of Labor Economics*, 32(2): 337–360.

- Perronnin, M., A. Pierre, and T. Rochereau.** 2012. “L’enquête Protection sociale complémentaire d’entreprise 2009.” *Les rapports de l’Irdes*, (554).
- Pertold, Filip.** 2019. “What if they take it all? Heterogeneous impact of sickness absence reform.” *Economics of Transition and Institutional Change*, 27(3): 647–671.
- Pettersson-Lidbom, Per, and Peter Skogman Thoursie.** 2013. “Temporary disability insurance and labor supply: evidence from a natural experiment.” *The Scandinavian Journal of Economics*, 115(2): 485–507.
- Pfarr, Klaus.** 2014. “femlogit-implementation of the multinomial logit model with fixed effects.” *Stata Journal*, 14(4): 847–862.
- Pollak, Catherine.** 2017. “The impact of a sick pay waiting period on sick leave patterns.” *The European Journal of Health Economics*, 18(1): 13–31.
- Puhani, Patrick A.** 2012. “The treatment effect, the cross difference, and the interaction term in nonlinear “difference-in-differences” models.” *Economics Letters*, 115(1): 85–87.
- Puhani, Patrick A, and Katja Sonderhof.** 2010. “The effects of a sick pay reform on absence and on health-related outcomes.” *Journal of Health Economics*, 29(2): 285–302.
- Sénat.** 2013. “Journée de carence non-respectée par le conseil général du Val-de-Marne. Question orale numéro 0280S de Mme Catherine Procaccia (Val-de-Marne - UMP).” <http://www.senat.fr/questions/base/2012/qSEQ12120280S.html>.
- Solon, Gary, Steven J Haider, and Jeffrey M Wooldridge.** 2015. “What are we weighting for?” *Journal of Human resources*, 50(2): 301–316.
- von Bieberstein, Frauke, and Jörg Schiller.** 2017. “Contract design and insurance fraud: an experimental investigation.” *Review of Managerial Science*, 1–26.
- Voss, Margaretha, Birgitta Floderus, and Finn Diderichsen.** 2001. “Changes in sickness absenteeism following the introduction of a qualifying day for sickness benefit-findings from Sweden Post.” *Scandinavian Journal of Public Health*, 29(3): 166–174.
- Ziebarth, Nicolas R, and Martin Karlsson.** 2010. “A natural experiment on sick pay cuts, sickness absence, and labor costs.” *Journal of Public Economics*, 94(11): 1108–1122.

**Ziebarth, Nicolas R, and Martin Karlsson.** 2014. “The effects of expanding the generosity of the statutory sickness insurance system.” *Journal of Applied Econometrics*, 29(2): 208–230.

---

# Appendices

## A Questionnaire on sick leave from the French Labour Force Survey

Individuals aged between 15 and 75 are first asked if during the reference week they have done at least one hour of paid work. If not, they are asked if they have nevertheless a job.

In the event that they have not worked but they have a job, they are asked why they have not worked in the reference week. One possible answer is "sick leave (including leave for a sick child) or work accident". Other possible answers include notably "paid days off, including days for the reduction of the working time ("RTT")", "maternity or paternity leave", "part time". If they are on sick leave or on leave related to a work accident, they are also asked what the total expected duration of the leave is.

In the event that they have worked during the reference week, they are asked if they have been "absent for sickness or a work accident", and if yes, how many days during the reference week.

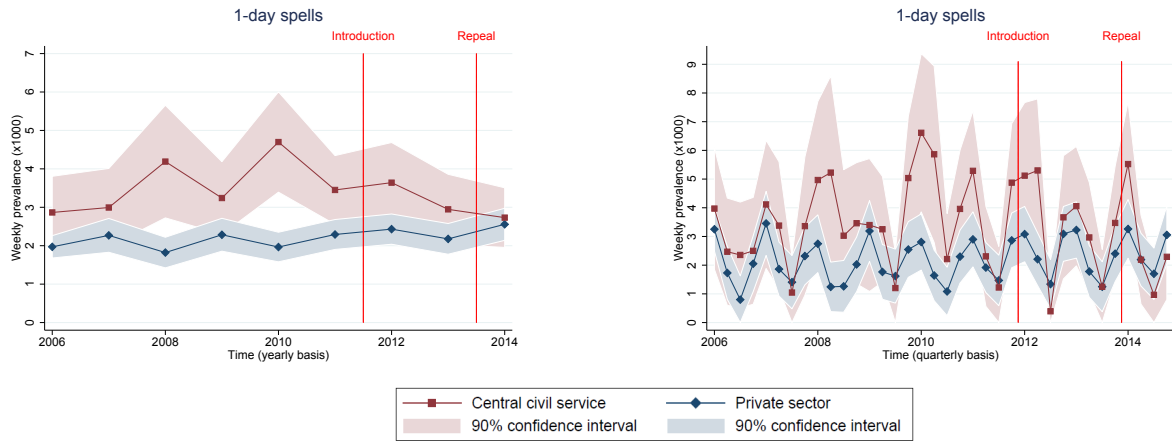
These two flows of questions related to sick leave provide two sick leave durations of different meaning. For those who have worked at last one hour and who have been on sick leave during the reference week (we will say they are on short-term sick leave), it is the duration of this leave during the reference week. Information may consequently be left-censored or right-censored. For those who have not worked at all during the reference week and who are on sick leave (we will say that they are on long-term sick leave), the information is neither left- or right-censored, but the duration is an estimated duration of the current sick leave at the time of the interview.

This part of the questionnaire remains almost unchanged for the years 2006-2014. The most notable change in 2013 concerns how the duration of the leave is measured (for those who have worked during the reference week). Up to 2012, it was measured in days or hours. After 2013, it is measured in days or half-days.

## B Prevalence of sick leave spells over time and sector, by duration category

**FIGURE B.I.**

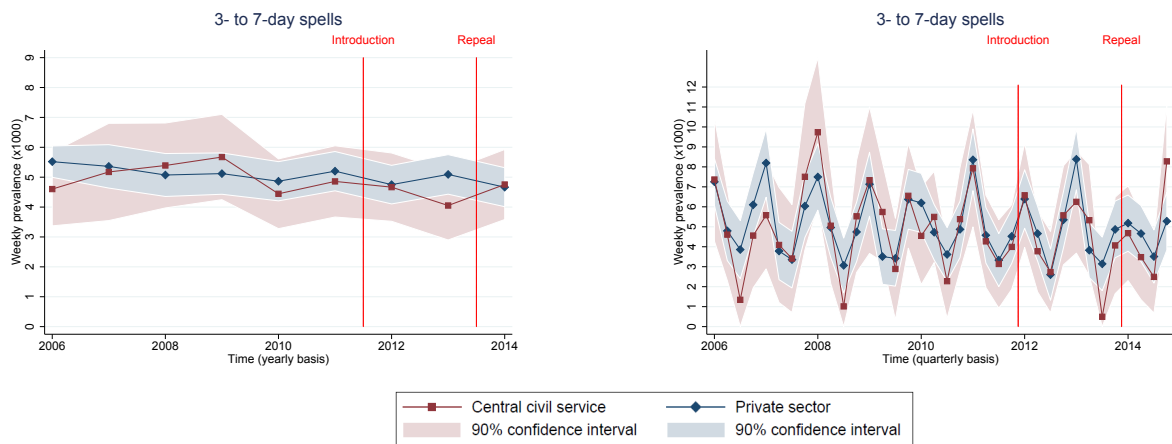
Prevalence of 1-day spells by sector, at a yearly (left) and quarterly (right) basis.



Source: French Labour Force Survey 2006-2014.

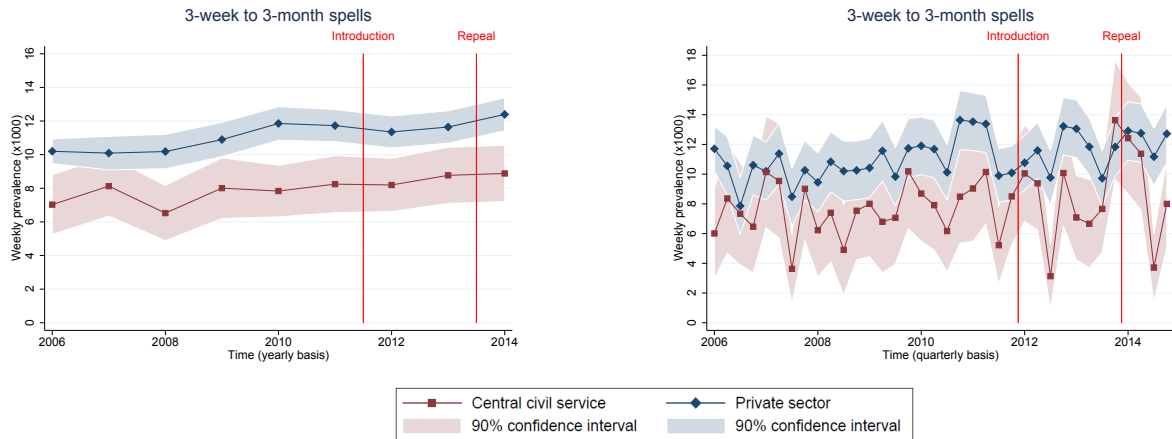
**FIGURE B.II.**

Prevalence of 3- to 7-day spells by sector, at a yearly (left) and quarterly (right) basis.



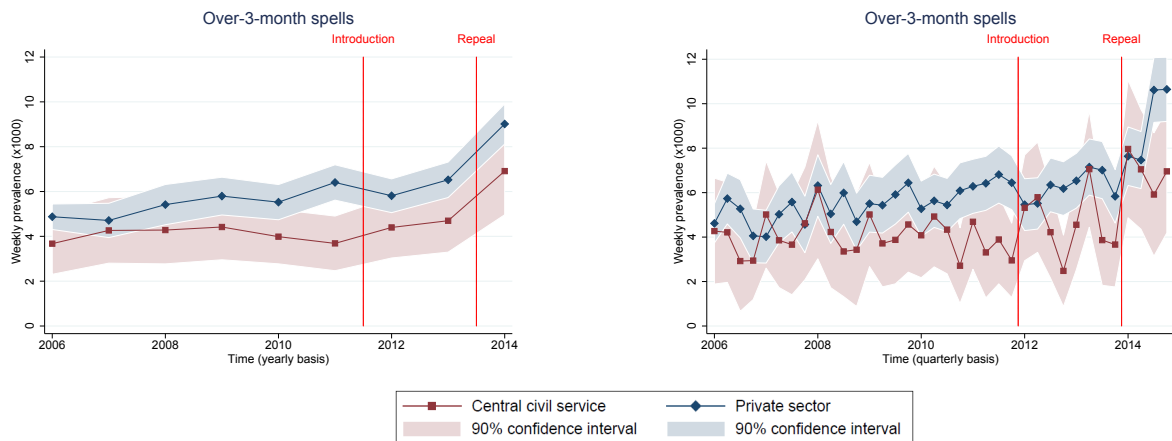
Source: French Labour Force Survey 2006-2014.

**FIGURE B.III.**  
Prevalence of 3-week to 3-month spells by sector,  
at a yearly (left) and quarterly (right) basis.



Source: French Labour Force Survey 2006-2014.

**FIGURE B.IV.**  
Prevalence of over-3-month spells by sector,  
at a yearly (left) and quarterly (right) basis.



Source: French Labour Force Survey 2006-2014.

# C Determinants of absenteeism

TABLE C.I. Determinants of absenteeism (prevalence  $\times$  1000), as shown by pooled OLS regressions

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
Women	0.850*** (0.145)	0.651*** (0.123)	0.941*** (0.212)	1.101*** (0.215)	2.977*** (0.384)	0.719* (0.404)	7.240*** (0.772)
Age	-0.0440*** (0.00723)	-0.0330*** (0.00610)	-0.0253** (0.0106)	0.00894 (0.0107)	0.218*** (0.0193)	0.393*** (0.0207)	0.518*** (0.0379)
Highly educated (degree level)	0.738*** (0.221)	-0.521*** (0.161)	-1.784*** (0.256)	-1.861*** (0.248)	-5.620*** (0.438)	-3.034*** (0.462)	-12.08*** (0.995)
Being in a couple	-0.437** (0.170)	-0.360** (0.149)	-1.109*** (0.252)	-0.386 (0.247)	-1.079** (0.431)	-1.304*** (0.444)	-4.675*** (0.833)
Having a child under 6	0.773*** (0.215)	0.319* (0.179)	0.750*** (0.280)	-0.00464 (0.273)	1.340*** (0.459)	1.082** (0.442)	4.259*** (0.868)
Wage (thousand euros per month)	-0.106** (0.0502)	-0.0214 (0.0369)	-0.450*** (0.0791)	-0.562*** (0.0883)	-1.917*** (0.240)	-2.199*** (0.267)	-5.256*** (0.611)
Teachers	1.339*** (0.457)	1.185*** (0.350)	-0.464 (0.492)	-0.918** (0.437)	1.019 (0.772)	0.460 (0.825)	2.622* (1.583)
Working days per week	-0.0896 (0.0950)	-0.0176 (0.0708)	0.474*** (0.125)	0.514*** (0.134)	1.210*** (0.248)	1.533*** (0.260)	3.624*** (0.472)
Central civil service	0.316 (0.296)	0.187 (0.215)	0.711* (0.379)	-0.0503 (0.355)	-2.610*** (0.559)	-2.098*** (0.583)	-3.546*** (1.117)
Year 2010	ref	ref	ref	ref	ref	ref	ref
Year 2011	0.114 (0.213)	0.160 (0.188)	0.326 (0.313)	-0.128 (0.303)	0.137 (0.498)	0.455 (0.465)	1.064 (0.880)
Year 2012	0.251 (0.217)	-0.0530 (0.181)	-0.0718 (0.303)	-0.0828 (0.303)	-0.0457 (0.520)	0.115 (0.488)	0.114 (0.936)
Year 2013	-0.00210 (0.219)	-0.00922 (0.189)	0.280 (0.316)	0.361 (0.320)	0.0275 (0.526)	1.033** (0.511)	1.689* (0.969)
Year 2014	0.232 (0.226)	0.193 (0.196)	-0.209 (0.308)	0.238 (0.325)	-0.132 (0.529)	2.420*** (0.553)	2.742*** (0.993)
Calendar quarter 1	ref	ref	ref	ref	ref	ref	ref
Calendar quarter 2	-1.106*** (0.200)	-1.216*** (0.185)	-2.307*** (0.287)	-0.563** (0.282)	0.0288 (0.421)	0.114 (0.278)	-5.050*** (0.655)
Calendar quarter 3	-1.949*** (0.188)	-1.713*** (0.172)	-3.658*** (0.271)	-1.550*** (0.265)	-2.246*** (0.408)	0.369 (0.316)	-10.75*** (0.653)
Calendar quarter 4	-0.484** (0.214)	-1.045*** (0.186)	-1.635*** (0.297)	-0.174 (0.285)	0.261 (0.419)	-0.0910 (0.302)	-3.168*** (0.674)
Constant	5.137*** (0.611)	4.182*** (0.469)	7.014*** (0.832)	3.941*** (0.846)	2.478* (1.458)	-10.36*** (1.591)	12.39*** (2.825)
Observations	588,955	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035	146,035
R <sup>2</sup>	0.00066	0.00043	0.00073	0.00051	0.00202	0.00327	0.00454

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## D Treatment effects for the 8 quarters of implementation of the policy

**TABLE D.I.** Treatment effects on the prevalence of spells ( $\times 1000$ ), for the 8 quarters of implementation of the policy

	Spell category						All spells
	1-day	2-day	3- to 7-d.	1- to 3-w.	3-w. to 3-m.	Over-3-m.	
T $\times$ 2012 Q1	-0.302 (1.909)	-2.167* (1.275)	1.861 (1.947)	2.392 (1.691)	0.357 (2.470)	0.657 (1.650)	2.797 (4.336)
T $\times$ 2012 Q2	1.372 (1.902)	-2.006* (1.153)	-0.152 (1.714)	2.725 (1.880)	-0.958 (2.587)	1.873 (1.838)	2.853 (4.471)
T $\times$ 2012 Q3	-1.138 (1.039)	-2.599*** (0.924)	1.492 (1.531)	0.829 (1.394)	-1.120 (1.969)	2.197 (1.891)	-0.340 (3.508)
T $\times$ 2012 Q4	-0.366 (1.728)	-1.916* (1.085)	0.985 (1.938)	-0.632 (1.507)	-0.0641 (2.799)	1.882 (1.733)	-0.112 (4.423)
T $\times$ 2013 Q1	-1.943 (1.650)	-2.650* (1.392)	-0.283 (2.068)	4.689** (1.956)	-4.004 (2.739)	-0.446 (1.990)	-4.637 (4.646)
T $\times$ 2013 Q2	-1.266 (1.442)	-0.757 (1.207)	2.089 (1.986)	3.436* (1.850)	-3.556 (2.755)	0.530 (2.430)	0.476 (4.682)
T $\times$ 2013 Q3	0.346 (1.203)	-2.660*** (0.886)	-0.805 (1.269)	2.036 (1.570)	3.616 (2.472)	0.566 (2.084)	3.100 (3.968)
T $\times$ 2013 Q4	0.490 (1.540)	-0.828 (1.120)	-0.625 (1.904)	0.363 (1.668)	4.373 (2.988)	1.230 (1.782)	5.003 (4.413)
Observations	588,955	588,955	588,955	588,955	588,955	588,955	588,955
Individuals	146,035	146,035	146,035	146,035	146,035	146,035	146,035
$R^2$	0.00163	0.00188	0.00197	0.00114	0.00276	0.00408	0.00400

Linear regressions with individual fixed-effects.

Cluster-robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$





## Chapitre 2

# Incitations par les prix, monitoring et pratiques médicales des hôpitaux

**Résumé :**

Dans les systèmes de tarification à l'activité, on ignore encore dans quelle mesure les hôpitaux réagissent aux incitations ciblées. Nous évaluons si une incitation tarifaire délibérée et une politique de mise sous accord préalable ont réussi à augmenter les taux de chirurgie ambulatoire des procédures qu'elles visaient. Utilisant les informations sur les séjours dans les hôpitaux français sur la période 2006-2013, nous tenons compte de la sélection des procédures dans les politiques en utilisant les perspectives médicales des taux de chirurgie ambulatoires associées aux Groupes Homogènes de Malades, ainsi que la méthode du contrôle synthétique. Au lancement de l'une ou l'autre de ces politiques, des effets positifs et significatifs sont souvent constatés dans le secteur public. Aucun effet négatif n'est généralement constaté. Dans le secteur privé, il y a peu d'effets significatifs, sauf lorsque les deux politiques ont été combinées. Dans les deux secteurs, des effets sont rarement constatés lors de nouvelles extensions de l'une ou l'autre des politiques.

Ce chapitre a été co-écrit avec Engin Yilmaz (Drees).

\* \* \*

---

Can price incentives or the threat of monitoring  
induce hospitals to change their medical  
practices ? Evidence from two French policies.<sup>§</sup>

Alexandre Cazenave-Lacroutz ¶      Engin Yilmaz ||

**Abstract**

It remains unclear to what extent hospitals respond to targeted incentives in activity-based financing systems. In this paper, we assess if a price incentive and a distinct monitoring policy succeeded in increasing the rates of same-day surgery of selected procedures. Using discharge records from all French hospitals over the 2006-2013 period, we address selection issues by combining the synthetic control method and prospects of same-day surgery rates provided by medical learned societies. At the launch of either policy, positive and significant effects are often found in the public sector. Generally, no downside effects are found. In the private sector, there are few significant effects, except when both policies are combined. In both sectors, effects are rarely found in further extensions of either policy.

---

<sup>§</sup>We thank for their comments and suggestions Christian Agethen, Didier Blanchet, Françoise Bourgoïn, Eve Caroli, Philippe Choné, Andrew Clark, Jean Baptiste Combes, Brigitte Dormont, Malik Koubi, Christophe Le Guéhennec, Jérémy L'Hour, Clarissa Mang, Alice Halys, Hélène Hartmann, Clément Rallet, Lise Rochaix, Sébastien Roux, Fabien Toutlemonde, Lionel Wilner, Joachim Winter and seminar participants of the *European Health Economics Association 2018 conference*.

¶Université Paris Dauphine - Insee - Crest; alexandre.cazenave-lacroutz at polytechnique.org

||Drees; engin.yilmaz at sante.gouv.fr.

## 1 Introduction

While improving care quality is a priority of most developed countries, there is still a lot to learn about "*which quality-improvement strategies can help deliver the best care at the least cost*" (OECD, 2017). There are examples where quality monitoring has turned out to be effective, but how it is implemented is highly heterogeneous. Since most developed countries have finished adopting activity-based financing (Busse et al., 2013), prices, that is tariffs, could now be tuned to influence the mix of treatments provided (Januleviciute, Askildsen, Kaarboe, Siciliani, & Sutton, 2016), and thereafter to improve quality.

Yet so far evidence is mixed regarding the impact of tariffs on actual activity<sup>1</sup> within an activity-based financing framework. There is evidence that hospitals may respond to the tariff associated with each Diagnosis Related Group (DRG), i.e. the categories of similar patients for which hospitals receive the same lump sum.<sup>2</sup> But these findings are mixed: in some cases, they show no real impact on volume or quantity as found by Dafny (2005) or Shin (2018) in the US ; in other cases they show a positive elasticity of volumes to tariffs as Januleviciute et al. (2016) in Norway and Verzulli et al. (2016) in Italy, or a negative elasticity of

---

<sup>1</sup>Some articles have highlighted the presence of upcoding, that is the (administrative) shifting of a stay from one Diagnosis Related Groups (DRG) to a higher profitable one, when the classification is refined by age (Dafny, 2005), by birth weight (Jürges & Köberlein, 2015; Shigeoka & Fushimi, 2014) or by location (Shin, 2018). This directly affects the hospital income, but not its activity.

<sup>2</sup>Such evidence used to be mainly based on US studies on specific DRGs, like those by Newhouse (1989), Cutler (1995), Yip (1998), Gilman (2000), Lindrooth, Bazzoli, and Clement (2007), Heaton and Helland (2009) and Liang (2015). That of Papanicolas and McGuire (2015) stands out, as it showed that hip replacement activity was affected in England after the introduction of differential reimbursement levels for uncemented and cemented hip replacements, as compared to Scotland where no financial incentive existed. The more profitable activity increased following a financial incentive, despite being of lower quality. Recently, some articles have used quasi-natural experiments to study the impacts of tariffs on the volume of several DRGs (Dafny, 2005; Januleviciute et al., 2016; Salm & Wuebker, 2015; Shin, 2018; Verzulli, Fiorentini, Lippi Bruni, & Ugolini, 2016).

volume to tariffs as in Salm and Wuebker (2015) in Germany. Some even find that this impact can depend on the DRGs (as Januleviciute et al. (2016); Verzulli et al. (2016)). Some find that the reaction depends on the hospital's status (Dafny, 2005), while others do not (Verzulli et al., 2016). More research is therefore needed to understand how to best shape public policies accordingly, and whether it is possible at all. As such, this article assesses the main impacts of two policies aiming at steering hospital activity: one is based on tariffs and one is based primarily on monitoring.

The two studied policies both generate strong incentives towards a common policy goal, perceived as quality-enhancing: the development of same-day surgery. The first policy is a financial incentive implemented by the successive French health ministers, playing on their ability to refine (or "not to refine") DRG tariffs (Hafsteinsdottir & Siciliani, 2010).<sup>3</sup> Originally, stays with same-day surgery were grouped in a different DRG from the corresponding stays with full hospitalisation, and they were assigned a lower tariff. In 2009, for some DRGs the same tariff was assigned to same-day stays and inpatient stays: increasing the relative tariff of same-day stays (compared to inpatient stays) to such a "single tariff" makes the former unambiguously more attractive financially. Other DRGs were given single tariffs in 2012 and 2013. The so-called "single-tariff" policy was then generalised in 2014. The implementation of these single tariffs was an exogenous, proactive and well-publicised tariff change designed to change clinical practice. It may lead (or not) to an underprovision of full hospitalisation (Hafsteinsdottir & Siciliani, 2010). It is reminiscent of the concomitant English Best Practice Tariffs (BPT) (Gershlick, 2016), the evidence of whose effects is still mixed in the very few DRGs

---

<sup>3</sup>This latter article derives under some hypotheses when refining should be applied from a welfare perspective.

that have been studied<sup>4</sup>, but which has dramatically increased same-day cholecystectomy thanks to an increase in the same-day tariff above the full hospitalisation tariff (McDonald et al. (2012), Allen et al. (2016)).

The second studied policy, the "prior approval policy", was based on constraint and concomitantly implemented by another stakeholder, the French National Health Insurance (the CNAM for *Caisse Nationale d'Assurance Maladie*). It targeted some acts, that were not necessarily concerned by the single tariff policy. The CNAM could decide to monitor some hospitals with an under-average practice of same-day surgery for these acts, for a few months at a time. In that case, selected hospitals needed to request the prior approval of the CNAM before performing any inpatient surgery. This is similar to certain local measures that accompanied the English BPT incentive for same-day cholecystectomy, such as default admission for any patient to same-day surgery rather than conventional surgery (ANAP & HAS, 2014). Unlike the case of the English BPT, this "prior approval policy" was another well-defined national policy which is well documented and did not necessarily concern the same stays.

This unique, quasi-experimental setting in France enables us to study, in a single framework, the impact of two different policies on a large number of DRGs. We aim at providing more evidence on whether payers can really use proactive tariffs

---

<sup>4</sup>The BPT policy has been introduced since 2010 into a growing number of practices: stroke, hip fracture, same-day cholecystectomy and so on. The English policy-maker aimed at delivering the best clinical practice through adequate reimbursement of high-quality care, but evidence of its effects is mixed. On the one hand, the BPT for stroke seems to have had no impact on the available national quality and outcome indicators (McDonald et al., 2012). On the other hand, the BPT for hip fracture may have increased the proportion of patients receiving surgery within 48 hours and reduced mortality (McDonald et al. (2012), O'Neill, Kreif, Grieve, Sutton, and Sekhon (2016)) and the BPT for same-day cholecystectomy may have had immediate and sizeable effects (McDonald et al. (2012), Allen, Fichera, and Sutton (2016)). Note that O'Neill et al. (2016) show that McDonald et al. (2012)'s initial results on the BPT for hip fractures are altered when using alternative methods to Difference in Differences, such as synthetic control analysis.

or the threat of monitoring to induce hospitals to change their medical practices.

Much more specifically, our setting to study the impact of tariffs on the practice of same-day surgery is somewhat different from previous studies. By studying a large number of DRGs, we go beyond the case-study performed by Allen et al. (2016) regarding one British BPT that promoted the development of same-day practice (see above). Moreover, the financial incentive at play is lower in our case, as the same-day practice in Allen et al. (2016) was provided a much higher tariff than the alternatives. Our setting also contrasts with the study by Januleviciute et al. (2016) where they leverage the yearly changes in tariffs in Norway to notably establish (among other results) that "*a higher difference in price between DRGs with an overnight stay and without (i.e. day case) for a given diagnosis [does not] reduces the proportion of patients admitted as day cases*".<sup>5</sup> Indeed, in the Norwegian context, the tariff between a same-day stay and a full hospitalisation was systematically different ; furthermore there were almost no change in the tariffs for same-day stays. Unlike them, and like Allen et al. (2016), we also devote a great care to the methodological choice of our controls (due to the proactive nature of the studied policies that may entail selection effects).

In addition, such research may more generally offer a better understanding of how hospitals operate. Although many theoretical models consider it as a single entity that necessarily responds to financial incentives (see for instance Siciliani (2009)), hospitals involve stakeholders whose interests and information availability may vary (e.g. doctors vs. the hospital owner or hospital management), and they may be worth modelling separately (Zhang, Wernz, and Slonim (2016), Boudway, Marchand, and Sato (2004)). Moreover, recent empirical evidence in the US

---

<sup>5</sup>More generally, Januleviciute et al. (2016) find no response of surgical DRGs to prices, in contrast with the findings of Verzulli et al. (2016) in Italy.



(Sacarny, 2018) and in Hungary (Kiss, Kiss, & Váradi, 2019) show that the resulting agency problems can prevent financial incentives targeting hospitals from being effective. Observing no effect from the two policies under study (most particularly observing no effect from a strong financial incentive policy) might be an indication that such agency problems are particularly serious in France. This could occur in the context where the financial incentive policy is implemented, as this policy does not target doctors. On the one hand, surgeons' wages in French public and most private non-for-profit hospitals<sup>6</sup> are fixed nationally and primarily based on their length of service. Their wages do not depend on the performance of the hospital, and even less on the relative tariffs under scrutiny. On the other hand, surgeons of the for-profit private hospitals are paid based on the acts they perform, and not based on the tariff under scrutiny (which are paid to the clinics). What is more, surgeons' reluctance to perform same-day surgery is viewed as one of the main obstacles to the development of same-day surgery in France (Cash, Cash, & Dupillet, 2011).

To avoid any selection bias related to the choice of DRGs in either policy, for each incentivized DRG in 2008-2009-2012, we seek to create a synthetic control of its same-day surgery rate based on the rates of same-day surgery of unaffected DRGs ; this synthetic control is defined in order to mimic both the past changes in the incentivized DRG and its prospect regarding the mid-term rate of same-day surgery, as provided by learned medical societies. When we find that a significant increase in the practice of same-day surgery has been caused by either of the two policies, we further analyse other outcomes.

---

<sup>6</sup>Together they account for 60% of surgical stays in France. Later we call "public hospitals" the public hospital and the private non-for-profit hospitals that are paid by the National Health Insurance on the same basis (most private non-for-profit hospitals), and "private hospitals" the private for-profit hospitals and the few remaining private non-for-profit hospitals.

Effects differ between the public sector and the private sector, and depending on when the policy was applied. When the policies are first applied, positive and significant effects on rates of same-day surgery are often found in the public sector, also when either the single tariff policy or the monitoring policy is applied in a loose sense. In the private sector, significant effects are less often observed, except when both policies are combined. For both sectors, effects are rarely observed in further extensions of either policy. When there is an increase in a same-day surgery rate due to the policy, we generally do not find downside effects such as an increase in re-hospitalisations (proxying a decrease in quality) or in the total number of stays.

The remainder of the paper is organised as follows. Section 2 details the two policies. Section 3 presents the data and how we are able to study the two policies. Section 4 outlines the empirical strategy and Section 5 details the results. Section 6 discusses them.

## **2 The single tariff policy and the prior approval policy**

### **2.1 The practice of same-day surgery in France**

Since at least the early 2000s, French public authorities have been promoting same-day surgery, that is, operating procedures allowing the patient to leave on the day of the intervention. It is seen as a lever for improving the quality and safety of patient care, since it would reduce the risk of nosocomial infections; and would also improve the comfort of patients, in particular by allowing them to return home ear-

lier. In addition to this public health and quality objective, this form of care also aims at containing health care expenditure, since same-day surgery is less costly, even when post-hospitalisation costs are taken into account (ANAP & HAS, 2014).

France's delay in this area was documented in the late 1990s (Mahieu & Raffy-Pihan, 1998). Policy-makers then adopted different approaches to promote the development of same-day surgery. Their first measures were initially regulatory, with the abolishment of certain constraints in 2003<sup>7</sup>, but also pedagogical, with awareness-raising campaigns and dedicated training programs targeting practitioners from 2001 onwards. Since 2008<sup>8</sup>, action has gradually shifted towards two national policies: the single tariff policy, designed by the Ministry of Health, and the prior approval policy, implemented by the National Health Insurance.

## 2.2 The single tariff policy

The studied tariff incentive, the single tariff policy, consists in applying an identical tariff for same-day surgery and conventional care of low severity.<sup>9</sup> As French DRGs are refined according to the severity level, this is otherwise not the case<sup>10</sup>: same-day surgery usually receives a lower tariff. Such changes may occur on the

---

<sup>7</sup>Since 1999, a hospital had to drop 1 to 2 beds in other services in order to create a new place in a same-day surgery service. This obligation was lifted by Ordinance no. 2003-850 of 4 September 2003. See ANAP and HAS (2012), p84.

<sup>8</sup>Regulatory amendments have also eased certain technical operating conditions in 2012. Since 2013, organisational recommendations and medico-economic software tools have also been proposed by two health agencies. Those changes are beyond the scope of this paper.

<sup>9</sup>Severity of the stay depends on eventual pathologies that are associated with the main diagnosis, on patient age, on her eventual death at the end of the stay, and on the stay length (Hafsteinsdottir & Siciliani, 2010). There are typically 4 levels of severity (going from Level 1 for low severity to Level 4 for very severe cases), plus one additional level for same-day stays (in that case the additional level is called "J") or for short stays (in that case the additional level is called "T"). For the ease of exposition, we consider as a single DRG all the corresponding levels of severity.

<sup>10</sup>As a caveat, in the French DRG classification before 2009, same-day stays received the same tariffs as 1-day stays, but a lower tariff than stays of greater length.

1<sup>st</sup> of March, since all tariffs are publicly modified at that date for the year to come. Note also that tariffs for public hospitals and for private-for-profit hospitals are separate.<sup>11</sup>

When a single tariff is implemented, hospitals are therefore encouraged to increase their rate of same-day surgery. In addition, since the cost of a same-day stay is lower than that of the same stay with complete hospitalisation, with a single fee, the margin of the establishment is then necessarily higher with a same-day stay.

These single tariffs have been progressively deployed in successive waves (see ANAP and HAS (2014); Bert et al. (2014); Cash et al. (2011)). For example, in the public sector, in 2009, the policy applied to 17 of the 223 surgical DRGs of the studied field. It was extended in 2012 to 38 DRGs, then in 2013 to 43 DRGs.<sup>12</sup> It was finally generalised in 2014 to all the DRGs for which same-day surgery was possible. An increasing percentage of stays has been affected by the policy (see Table 1).

### 2.3 The prior approval policy

At the same time, the National Health Insurance also sought to promote same-day surgery through a procedure of prior approval of some lists of acts. This policy consists in defining lists of acts that can be routinely performed on a same-day basis and in publishing these lists. At regional level, the National Health Insurance can then choose hospitals among those below their regional average in terms

---

<sup>11</sup>For those reasons, we consider the "financial years" from March to February, and we study the hospitals with the tariffs of the public sector (mainly public hospital and most private-non-for-profit hospitals) separately from those with the tariffs of the private-for-profit sector (mainly private hospital and some private-non-profit hospitals).

<sup>12</sup>In specific cases, when the single tariff resulted from the drop of deduction for short-stay outliers without the creation of a specific same-day severity level for the DRG, the presence of the single tariff could differ between the public sector and the private sector. For instance, DRG 08C28 has thus had a single tariff in 2009 in the private sector, but not in the public sector.

Table 1: **Scaling up of the single tariff policy**

Financial year	DRGs with a single tariff		Percentage of stays with a single tariff	
	private sector	public sector	private sector	public sector
2006	0	0	0 %	0 %
2007	0	0	0 %	0 %
2008	0	0	0 %	0 %
2009	20	17	41.3 %	23.9 %
2010	20	17	41.4 %	24.7 %
2011	20	18	41.4 %	24.9 %
2012	41	38	54.2 %	36.9 %
2013	46	43	58.1 %	40.5 %
2014	98	105	96.8 %	95.6 %

**Note:** Financial year 2008 corresponds to the period from 1<sup>st</sup> March 2008 to 1<sup>st</sup> March 2009 not included. In this table, months of January and February of calendar year (A+1) are not included in tariff year A because they were not available for the 2014 financial year.

**Reading:** In 2009, 23.9% of stays in the field of study (see below) in the public sector benefited from a single tariff.

**Field of study:** Surgical stays (save surgical operations on pathological pregnancies and births) in France from 2006 to 2014.

**Source:** ATIH, PMSI-MCO (version v11e of the French DRG classification), treatment by the authors.

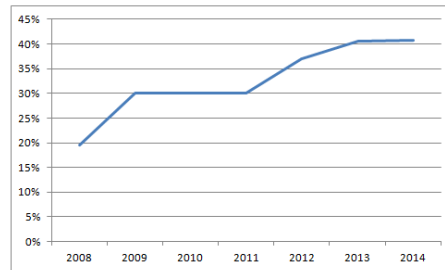
of same-day surgery. For a maximum period of 6 months (unless renewed), these hospitals must obtain a prior agreement from the National Health Insurance for stays in full hospitalisation if they want these stays to be priced at the tariff of the full hospitalisation. In practice, whenever the hospital's doctor thinks that a patient cannot be treated on a same-day basis for one of the monitored acts<sup>13</sup>, she has to prove it to a doctor of the National Health Insurance.

This binding procedure was scaled up at the same time as the single tariff policy: it concerned 5 lists of acts in 2008, 17 in 2009, 38 in 2012 and 43 in 2013.<sup>14</sup>

<sup>13</sup>and if there is no emergency.

<sup>14</sup>The selected lists are available at:<https://www.ameli.fr/medecin/exercice-liberal/prescription-prise-charge/accord-prealable/accord-prealable-chirurgie-ambulatoire>.

Admittedly, the procedure effectively concerned a limited number of hospitals (less than 200 in 2012 (Bert et al., 2014)) but it may also have affected practices of hospitals not concerned by the monitoring procedure themselves. Indeed, it is likely that no hospital wanted to be subject to such a binding procedure; moreover, this procedure brings a clear information, at national level, about which acts can be routinely practiced in same-day stay. Thus, we consider that a stay was subject to the prior approval policy if it includes one of the acts concerned by this policy (see Part 3 for more details), without taking into account whether or not the hospital has actually been placed under a prior control procedure by the National Health Insurance. This is what we later call "loose prior approval", in contrast to "strict prior approval" when the hospital is actually subjected to prior approval for a given act (see graph 2). A focus on loose prior approval rather than strict prior approval ensures that potentially affected stays are not considered as unaffected: this conservative approach is linked to our evaluation strategy (we compare "incentivized" DRGs with unaffected DRGs) and would be fully justified *ex-post* if we obtain effects of loose prior approval that are too large to result from strict prior approval. Under this loose definition, a large and growing proportion of stays were concerned by the prior approval policy (see Figure 1).

Figure 1: **Scaling up of the prior approval policy**

**Note:** Financial year 2008 corresponds to the period from 1<sup>st</sup> March 2008 to 1<sup>st</sup> March 2009 not included. In this table, months of January and February of calendar year 2015 are not included in tariff year 2014.

**Lecture:** In 2008, 22% of stays of the field of study (see below) might theoretically be concerned by a procedure of prior approval policy. We consider that they are impacted by the prior approval policy.

**Field of study:** Surgical stays (save surgical operations on pathological pregnancies and births) in France from 2006 to 2014.

**Source:** ATIH, PMSI-MCO (version v11e of the French DRG classification), treatment by the authors.

## 3 Data and the studied treatments

### 3.1 Data

Our data cover all the hospital stays in France for the period from March 2006 to February 2014. Many variables are present as this dataset is used to determine to which DRG each stay belongs. It contains patient information on the calendar month of the start of the stay, the length of the stay, the DRG that would have been used for invoicing during the 2013 financial year as well as the DRG that was effectively used for the invoicing, the patient's gender and age, the number of comorbidities of the stay, the number of acts performed, and so on. As it includes all hospitalisations in France, we are able to compute readmission indicators (see Yilmaz and Vuagnat (2015a)). Similarly to Allen et al. (2016), we have restricted the scope of our study to surgery, excluding surgical operations on pathological pregnancies and births. Table 10 of Annex A reports the main features of our database.

We aggregated those stays across DRGs (without taking into account the severity level), at national level, separating those under public sector tariffs and those under private-for-profit tariffs and at monthly frequency. We use a homogeneous DRG-grouping, consisting in the DRG grouping for financial year 2013.<sup>15</sup> This enables us to bypass minor changes introduced every year into the classification (with some DRGs being refined or mixed, for instance).

We focus our study on the same-day surgery rate , as this is the target of the two policies. Same-day surgery rate is the percentage of stays without a night at the hospital, out of the total number of stays within the DRG. But we also study other variables in a complementary section, as robustness tests or since they can also be important from a policy perspective. First, an increase in same-day surgery rate should mechanically come from a substitution effect, due to a decrease in the rate of stays of other severity (especially stays of severity 1). The percentages of stays of higher severity and of higher lengths (1-night stays, 2-night and 3-night stays) are studied in order to confirm the diagnosis established with the same-day surgery rate . Second, even though same-day surgery is thought to convey a better quality of care by French governments, evidence is scarce (Allen et al., 2016; ANAP & HAS, 2014). A better proxy for the quality of care is readmission rates. We study the readmission rate within 1 week and within 1 month for each DRG.

Thirdly, the total number of stays of the DRG is studied to check whether the "induced demand" hypothesis might be at work, that is whether the measure may have encouraged hospitals to treat patients they would not have treated (at all) without the financial incentives.<sup>16</sup>

---

<sup>15</sup>Such a homogeneous DRG-grouping has already been used in France by Milcent (2020).

<sup>16</sup>This was observed for instance by Verzulli et al. (2016) in Italy after an increase in tariffs, and by Januleviciute et al. (2016) after a decrease in tariffs.



We also collected complementary data: tariffs were publicly available for each year on the website of the Technical Agency for Hospital Information (ATIH), and the lists of acts subject to the prior approval policy were found on the National Health Insurance website. Finally, between 2010 and 2013, seven medical learned societies<sup>17</sup> established prospects of rates of same-day surgery deemed to be potentially achievable in the "medium-term" (or "by 2016") for the DRGs of their speciality (in the rest of the chapter: the "*same-day prospects*").<sup>18,19</sup> These same-day prospects have a lower-bound rate and a higher-bound rate. We collected these same-day prospects for 153 DRGs over the 223 DRG (69%) in our field of study.

### 3.2 The two treatments

This section defines what the two treatments for a  $DRG \times month \times sector$  unit are. Indeed, section 2.2 and section 2.3 have defined what we call treatments for a hospital stay, but the link is not immediate for two different reasons.

Firstly, the single tariff policy is applied according to the DRG-classification whereas the prior approval policy is applied according to the acts classification. We group stays along the DRG-classification mainly because we have at our disposal same-day prospects by DRG - and not by act (see above). As there is no correspondence between these classifications, a DRG can encompass across its stays

---

<sup>17</sup>Their specialities are: gynecological surgery, ENT, stomatology, orthopedics, ophthalmology, urology and vascular and digestive surgery.

<sup>18</sup>When a DRG is in the scope of two learned societies, we have computed the average of each rate.

<sup>19</sup>Access to these data required our commitment not to disseminate them, because some of the learned societies made their confidential nature a prerequisite for their collaboration with the department of the Health Ministry which gave us access to these data.

many different acts, some that were concerned by the prior approval policy, and some that were not. Hence, we have to make a choice regarding which DRGs can be considered as unaffected by the prior approval policy. We have calculated for each DRG and each financial year the percentage of stays for which at least one act concerned by the prior approval policy is included. We consider that a DRG potentially has the prior approval treatment when that percentage is greater than 10 %. This rather low threshold limits the risk that DRGs affected by the policy might be considered in the control group.

Secondly, we also have to make a choice regarding which DRGs were affected by the single tariff policy. There would be no such choice to make if, in year Y, we were using the DRG classification of year Y, as each DRG would have a single tariff or not. But we have to make such a choice, as across all the years we use the 2013 DRG grouping (see Section 3.1). For instance, if a new DRG is created from stays that benefited from a single tariff and other that did not, it is not easy to determine which treatment status should be attributed to this DRG for the previous years.<sup>20</sup> In practice, before aggregation, each stay was identified as having actually been affected by the single tariff policy or not ; we then considered that the 2013-DRG was not affected by the single tariff policy [respectively: was affected] when the percentage of stays actually concerned that year was less than 5 % [resp.: greater than 95 %]. There are very few cases left.<sup>21</sup>

---

<sup>20</sup>Assuming there was a single tariff for its stays is only partly true and leads to underestimating the impact of the introduction of a single tariff. Assuming it was not the case and placing it among the control DRGs may underestimate the impact of the incentive on the other incentivized DRGs - if the incentive has an effect.

<sup>21</sup>We list here the few left cases and how we proceed with them. For instance, since DRG 03C09 had 90 % of its stays with a single tariff from 2012 on, it was considered as affected by the policy. Similar choices were made in the private sector (respectively: in the public sector) for DRGs 02C10, 06C25, 08C04, 08C13, 08C56b, 12C04b and 21C06 (respectively: 02C12, 06C25, 08C13 and 12C04b). DRG 02C12 which had 50 % of its stays that were affected in 2009 in the private sector, then none in the following years, was considered as unaffected.

## 4 The Empirical strategy

### 4.1 Difficulties

Our aim is to assess the effect of both policies on rates of same-day surgery of incentivized DRGs. This identification is not trivial. Three points, for instance, make it unlikely that a direct comparison between incentivized DRGs and all unaffected DRGs would provide reliable estimates of these effects.

First and foremost, incentivized DRGs have been selected; they were not randomly picked. Hence, direct comparisons would not control for the evolution of rates of same-day surgery, which might for instance bias the results upward if selected units were already experiencing a high increase before the policy (which occasionally happens). It gives indeed a hint that it could have continued even without the policy (which is not observable).

Secondly, direct comparisons would similarly not control for any new medical advances that may suddenly ease the practice of same-day surgery for some DRGs. It might for instance bias the results upward if the selected units were chosen among units where a higher rate of same-day surgery is easily accessible, which is unobservable to the researcher.<sup>22</sup>

Thirdly, the few case studies on the British BPTs raise suspicions, for instance, that such financial incentive produce heterogeneous impacts, with incentives having no effect in some cases. Hence, it would be desirable to estimate a specific effect for each DRG.

---

<sup>22</sup>And we know that this has been attempted by the health policy-maker, as shown by ATIH (2014) which clarifies how DRGs were selected to benefit from the single tariff. Before 2011, these DRGs were chosen based on medical criteria: According to some experts, same-day surgery could partially or completely substitute complete hospitalisation for these DRGs. In 2012 however, a technical criterion was used: the average duration of stays should be less than 3 days, and there should be 2/3 of acts in common between stays of the DRG for the same-day surgery and the stays of the same DRG with severity level 1.

The following sections present our strategy to take these selection effects into account. In spirit, our strategy is inspired by Allen et al. (2016). But we systematise it through the use of the synthetic control procedure developed by Abadie, Diamond, and Hainmueller (2010). It also allows us to obtain an estimate per incentivized DRG. As a caveat, note that we will not be able to distinguish between what comes from the single tariff policy and what comes from the prior approval policy when both policies affect a DRG.

## **4.2 The Methodology**

For each incentivized DRG, we attempt to build a "credible" counterfactual rate of same-day surgery, based on the rates of same-day surgery of unaffected DRGs. For DRGs with a policy starting in 2009 (respectively 2012 ; 2008), we consider that DRGs are unaffected if no policy was applied to them until 2012 (respectively 2014 ; 2011), which enables us to analyse the effects of the policy in 2009, 2010 and 2011 (respectively in 2012 and 2013 ; in 2008, 2009 and 2010). For each incentivized DRG, we build a counterfactual rate of same-day surgery made up of a combination of the rates of same-day surgery of unaffected DRGs such that, first, pre-incentive counterfactual evolutions of same-day surgery rates should reproduce those of the studied DRGs, and, second, their same-day prospects should be similar. In practice, we used a version of the synthetic control procedure developed by Abadie, Diamond, and Hainmueller (2015), presented in more detail in Section 4.3.

To test the relevance of the created counterfactual same-day surgery rate (see Section 4.5), we check that each synthetic control reproduces the trajectory of same-day surgery rate of the incentivized DRG before the incentive, by consider-

ing both the average difference and the average slope of this difference over the period preceding the incentive. We make the final analysis only for the incentivized DRGs whose synthetic control seems to be a reasonable counterfactual. To further assess whether they are reasonable, we also present the differences in prospects between each incentivized DRG and its synthetic control and we implement temporal placebos, as suggested by Abadie et al. (2015).

### 4.3 Creation of a synthetic control rate of same-day surgery

Abadie and Gardeazabal (2003) and Abadie et al. (2010) developed a numerical procedure to establish a counterfactual unit reproducing the trajectory of the interest variable before the incentive in case the number of possible controls is limited. As reported by Abadie et al. (2015), this approach provides an explicit mechanism for selecting comparison units (here, the DRGs), even in the event of small samples. The synthetic control is built to approximate relevant characteristics (relative to the studied variable) of the studied unit before the implementation of a policy, *via* a convex combination of available units of control (see below). While initial applications of this statistical method were to estimate the effect of a policy on a single unit, Billmeier and Nannicini (2013) (and others) have also used it to evaluate the heterogeneous effects of a policy on several different units. It was also used among other statistical methods by O’Neill et al. (2016) to show that some results on the impact of an English Best Practice Tariff (BPT) were not robust due to a possible violation of the common trend assumption hypothesis.

Concretely, for each incentivized DRG for which we have mid-term same-day prospects, we calculate a counterfactual rate of same-day surgery from rates of same-day surgery of unaffected DRGs for which we also have same-day prospects.

To do this, we determine the convex combination of these DRGs that best approaches the two same-day low and high prospects of incentivized DRGs as well as the trajectory of same-day surgery rate of incentivized DRGs before the policy was applied to them. If this combination approximates the trajectory of incentivized DRGs before the start of the policy, our hypothesis is that it would also have approximated it in the absence of the policy. We technically formalize this approach below. It should be noted that, as in Allen et al. (2016), the unaffected DRGs at the origin of the synthetic control are not necessarily similar to the incentivized DRG in terms of medical content: for example, the synthetic control of the same-day surgery rate of an incentivized DRG treating eye surgery may potentially combine the changes in rates of same-day surgery of DRGs performing hand or stomach surgery. Indeed, only the possibilities of changes to the same-day surgery rate defined by the learned medical societies are used here as selection criteria.

Let  $i$  denote the DRGs, ranging from 1 to  $N$ . DRG  $i = 1$  is affected by a policy from the period  $T_d$  to the period  $T_f$  whereas the other DRGs are not affected. For all  $(i, T)$ ,  $Y_{i,T}$  is the observed rate of same-day surgery for DRG  $i$  at period  $T$ . For all  $i$ ,  $Z_i$  is a vector of some  $Y_{i,t}$ <sup>23</sup> and of the two lower-bound rate and higher-bound rate of same-day prospects. We consider the convex combinations

---

<sup>23</sup>Like Abadie et al. (2010), we use as explanatory variables past records of the dependent variables (in our case past rates of same-day surgery), and other covariates. Yet Kaul, Klößner, Pfeifer, and Schieler (2017) have shown that one cannot use both the entire trajectory of the variable of interest pre-incentive and other control variables, because then the latter are no longer considered by the optimisation procedure. Hence, we do not consider the whole trajectory, but only include the twelve observations preceding the policy applied and the monthly averages before it applied, for each month. This should make the synthetic control a good counterfactual regarding average value before the policy applied, trends immediate before the policy applied, and seasonal variations in the practice of same-day surgery. As explanatory variables, we also include both the lower-bound rate of same-day prospects according to the learned medical societies and the corresponding higher-bound rate (See section 3.1). This may help the synthetic control to be a good counterfactual regarding other characteristics that are unobservable in the past trajectories of rates of same-day surgery but that doctors (i.e. the learned medical societies) observe.

of DRGs  $i = 2..N$ :  $\sum_{i=2}^N w_i \cdot Z_i$  where  $w_i$  is the weight associated with DRG  $i$ :  $w_i \geq 0, \sum_{i=2}^N w_i = 1$ .

The weights  $(w_i^*)$  used to build the synthetic control are chosen so as to minimize the difference between the rates of same-day surgery of such convex combination and the studied DRG  $i = 1$ , before the incentive is applied. With  $V$  a positive-definite matrix that defines a distance, the synthetic control is:

$$\tilde{Z}_1 = \sum_{i=2}^N w_i^* \cdot Z_i \quad (1)$$

with:

$$(w_i^*)_{i=2..N} = \operatorname{argmin}_{(w_i)_{i=2..N}} (\| (Z_1 - \sum_{i=2}^N w_i \cdot Z_i)_{T \leq T_d-1} \|_V) \quad (2)$$

As in Abadie et al. (2010, 2015), one possibility is to choose  $V$  among all the diagonal and defined positive matrices and so as to make the optimisation on both  $V$  and the  $(w_i)$ :

$$(w_i^*)_{i=2..N} = \operatorname{argmin}_{(V, (w_i)_{i=2..N})} (\| (Z_1 - \sum_{i=2}^N w_i \cdot Z_i)_{T \leq T_d-1} \|_V) \quad (3)$$

Most notably for computational ease, we do not carry out the complete optimisation, that is, on both the matrix  $V$  and the vector of  $(w_i)$ . When we first consider the data-driven matrix supplied as a standard option by Abadie, Diamond, and Hainmueller (2014)'s Stata program,<sup>24</sup> both the lower-bound rate and the higher-bound rate of same-day prospects are given negligible weights.<sup>25</sup> To ensure a sufficient weight for the lower-bound and the higher-bound rates of same-day prospects, we have chosen the matrix  $V$  which attributes 25% of the weight to the two lower-bound and higher-bound rates of same-day prospects, and then

<sup>24</sup>How this matrix is computed is described in Abadie, Diamond, and Hainmueller (2011).

<sup>25</sup>These preliminary results are displayed in Annex D.

equal weight to the other variables.<sup>26</sup>

For any month  $m$ , the estimated effect is then:  $e_{1,m} = Y_{1,m} - Y_{CS,m}$ . In practice, to present in tables a single effect for a given period  $a$  (for instance one year), we also compute  $E_y$  the (simple) average of the estimated monthly effects:

$$E_y = \frac{1}{12} \sum_{m \in \text{year } y} (Y_{1,m} - Y_{CS,m}) \quad (4)$$

#### 4.4 Inference with a synthetic control estimator

The previous method yields an estimated effect per incentivized DRG and period, expressed in percentage points of rate of same-day surgery. If the optimisation succeeded, by definition it should be very close to zero before the policy applied. The optimisation no longer imposes this constraint after the policy starts to be applied, and any deviation after  $T_d$  may either signal an impact of the policy or (for instance) an idiosyncratic deviation between the incentivized DRG and its synthetic control. Hence, for the incentivized DRG we need to assess the significance of the "estimated effect" after the policy starts to be applied. To do so, a Fisher exact inference test, or "placebo spatial test", is applied: for each of the (N-1) unaffected DRGs, a placebo statistic is calculated by applying exactly the same procedure (by considering in the construction of its synthetic control the (N-2) other unaffected DRGs and the incentivized DRG). The estimated effect is reputed to be statistically significant if the statistic computed for the incentivized DRG is higher than the statistics of the placebo tests.

The statistic used in this inference procedure is not directly the estimated effect.

---

<sup>26</sup>This choice was arbitrary. What matters is that we can control *ex-post* that it was generally sufficient to enforce a good optimisation regarding same-day prospects. When we increase this weight to 25% for each of the two variables, results do not differ much (see Annex E).



Indeed, in practice, the optimisation procedure (i.e. the construction of synthetic controls) is not successful with some DRGs, especially those that are very different from all the others. While the difference with their synthetic control DRG should be almost zero before the start of the policy, it can be quite large beforehand for these outliers - as well as afterwards. A "large" gap after the incentive is not very informative if it was already "large" before the incentive. To penalise these cases where optimisation has failed, Abadie et al. (2010, 2015) recommend, for the inference, using the ratio of the mean squared prediction error (RMSPE) (see below). Firpo and Possebom (2018) confirm on simulated data that the resulting inference test is more powerful than other, more standard, tests. To determine the significance of the estimated effect, we therefore use an exact Fisher inference test on the RMSPE. We now detail how the RMSPE is constructed.

We denote  $RMSPE_{g,y}$  the RMSPE corresponding to DRG  $g$  and year  $y$  (with  $y \geq T_d$ ). It is the ratio of the sums of the squared prediction errors, between year  $y$  and the years before the policy started.

$$RMSPE_{g,y} = \frac{\frac{1}{12} \cdot \sum_{m \in \text{period } y} (e_{g,m})^2}{\frac{1}{T_d-1} \cdot \sum_{T=1}^{T_d-1} (e_{g,T})^2} \quad (5)$$

A high RMSPE indicates that the difference, after the incentive, between the same-day surgery rate of the incentivized DRG and the same-day surgery rate of its synthetic control is high compared to the same difference before the incentive. To assess the significance of the estimated effect on the incentivized DRG compared to the placebo estimated effects, we compare the estimated RMSPE. The DRGs are therefore ranked in descending order of RMSPE : we define the *p-value* of the estimated effect of the incentivized DRG as the ratio of the rank of the incentivized DRG to the number of considered DRGs ( $N$  here), expressed as a

percentage.

$$p - value_{g,y} = \frac{rank(RMSPE_{g,y})}{N} \times 100 \quad (6)$$

A p-value lower than 5% for the incentivized DRG signals that more than 95% of DRGs in the control group have a lower RMSPE than the incentivized DRG when the methodology is applied to them. Or in other words, that fewer than 5% of the placebo effects are stronger than those found for the incentivized DRG, which might validate the presence of an effect of the incentive (if our further robustness checks are validated). A p-value of more than 10% indicates that more than 10% of the placebo effects are stronger than those found for the incentivized DRG, which would lead us to conclude that the effect is not statistically significant.

## 4.5 Further restrictions

We check that the constructed synthetic control is a credible counterfactual of the incentivized DRG, in three ways.

Firstly, we check that the synthetic control reproduces the trajectory of rate of same-day surgery of the incentivized DRG before the incentive, with regard to the average level as well as the average slope of the difference between the two in the period preceding the incentive. There may be a significant gap, especially when the incentivized DRG is an extreme case compared to the control DRGs used to build its synthetic control. When the average slope of the same-day surgery rate of the incentivized DRG is greater than the average slope of the synthetic control, this also suggests that the selection problem has not been solved: even by choosing a synthetic control similar to the incentivized DRG, the growth of the incentivized

DRG is more dynamic than that of the synthetic control, and this may still be the case after application of the incentive, even in the absence of this incentive.

A case-by-case graphical analysis eliminates obvious cases where the optimisation failed. For any significant effect obtained after one year of incentive (respectively two years), we also check whether it is at least 3 times higher than the average pre-incentive difference, and at least 3 times higher (respectively 6 times higher for the effect at two years) than the average annual pre-incentive slope of the gap.<sup>27</sup> For any non-significant effect, we also check whether the mean pre-incentive deviation and the pre-incentive mean annual slope of the deviation are less than 5% of the pre-incentive mean value. Hence, we no longer study the DRGs whose constructed synthetic control fails to pass these restrictions.

Secondly, we perform and present "time placebo tests", as suggested by Abadie et al. (2015). To do so, we move backward by one year all the incentive dates and compute the corresponding estimates. When a significant effect is obtained in the year preceding the actual incentive year, this lowers our confidence that the selection problem has been solved: even by restricting the control group to "similar" DRGs, and even by reproducing the past trajectory of the same-day surgery rate, the growth rate of same-day surgery of the incentivized DRG is (immediately before the incentive) more dynamic than the one of its synthetic control even in the absence of incentive.

Thirdly, the differences in same-day prospects (as assessed by the learned medical societies) between the incentivized DRGs and their synthetic control are also reported. If the estimated effect after three years is of the order of magnitude

---

<sup>27</sup>The choice of these thresholds is arbitrary; defining thresholds in relation to the observed order of magnitude makes it possible to systematise this test on the various studied variables (rather than having to arbitrarily choose a threshold per DRG and per variable).

of these differences, it means that the differences in prospects between the incentivized DRGs and the unaffected DRGs have not been sufficiently considered; above all, it may give rise to concerns that the estimated effect does not reflect the causal impact of either policy, but just a natural increase due to higher achievable rates of same-day surgery. This is in general not the case in our current specification.

## 5 Results

### 5.1 Rates of same-day surgery for DRGs with the single tariff policy

Out of the 10 DRGs in the public sector with a single tariff starting in 2009 and for which we are able to build a reasonable synthetic control,<sup>28</sup> a positive and significant effect appears in 8 cases (see Table 2).<sup>29</sup> This occurs whether the single tariff is accompanied by a sizeable loose monitoring policy (in 4 cases<sup>30</sup> out of 4) or not (in 4 cases<sup>31</sup> out of 5). In those cases, the effect after three years is always much higher than the residual difference of same-day prospects between the studied DRG and its synthetic control.

Furthermore, of the 14 DRGs in the public sector with a single tariff starting in

---

<sup>28</sup>For those that we tried to study but for which we were not able to build a reasonable synthetic control, the graphs are however displayed in Annex C.

<sup>29</sup>The corresponding graphs for those incentivized or in 2009 are in Figure 2, and in Figure 6 of Annex B for those incentivized starting in 2012.

<sup>30</sup>This concerns: DRG 01C14: *Superficial nerve release with the exception of the median carpal tunnel*, DRG 13C12: *Dilation and curettage, conations for non-malignant conditions*, DRG 06C12: *Restorative procedures for inguinal and crural hernias, age greater than 17*, and DRG 09C07: *Local biopsies and excisions for nonmalignant breast conditions*.

<sup>31</sup>This concerns: DRG 03C21: *Interventions for protruding ears*, DRG 09C08: *Interventions on the anal and perianal region*, DRG 13C08: *Interventions on the vulva, vagina or cervix*, and DRG 09C13: *Interventions for anogenital condyloma*

2012, a positive and significant effect appears in only 2 cases.<sup>32</sup> Both also benefited from the monitoring policy.

Overall, this constitutes evidence that the single tariff policy had an effect most of the time within the public sector when first applied in 2009, whether it was accompanied by the monitoring policy or not. When applied in 2012 however, we do not find an effect in most cases.

Of the 16 DRGs in the private sector with a single tariff starting in 2009 and for which we are able to build a reasonable synthetic control, a positive and significant effect appears in 7 cases (see Table 3).<sup>33</sup> But one of these apparent effects has to be discarded as the single tariff was in fact removed from 2010 on.<sup>34</sup> Of the remaining 6 cases,<sup>35</sup> the single tariff is accompanied by a sizeable loose monitoring policy in 4 cases, and a weak loose monitoring policy in 2 cases.<sup>36</sup> A significant and positive effect thus appears in most cases when there was both the loose monitoring policy and the single tariff (i.e. in 6 cases out of 8).

Of the 14 DRGs in the private sector with a single tariff starting in 2012, a positive and significant effect appears in only 2 cases.<sup>37</sup>

Hence, there is no observable effect of the single-tariff policy alone in the private

---

<sup>32</sup>This concerns: DRG 08C58: *Arthroscopies of the shoulder* and DRG 07C14: *Cholecystectomies without exploration of the main bile duct except acute diseases*.

<sup>33</sup>The corresponding graphs for those incentivized in 2009 are in Figure 3, and in Figure 7 of Annex B for those incentivized starting in 2012.

<sup>34</sup>This concerns DRG DRG 12C05: *Interventions on testes for malignant tumors*

<sup>35</sup>This concerns: DRG 13C08: *Interventions on the vulva, vagina or cervix*, DRG 08C44: *Other interventions on the hand*, DRG 13C12: *Dilation and curettage, conations for non-malignant conditions*, DRG 01C14: *Superficial nerve release with the exception of the median carpal tunnel*, DRG 06C12: *Restorative procedures for inguinal and crural hernias, age greater than 17*, and DRG 09C07: *Local biopsies and excisions for nonmalignant breast conditions*.

<sup>36</sup>This is indeed the case for DRG 13C08 from 2010 on when a significant and positive effect starts occurring.

<sup>37</sup>This concerns: DRG 10C12: *Thyroid Interventions for Non-malignant Conditions* and DRG 08C58: *Arthroscopies of the shoulder*.

Table 2: Estimates in the public sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
08C12	2009 <sup>α</sup>	yes	no	18.2	2.96	1.51	3.67	3.76	-1.5	1.6	2.29
03C21	2009	yes	no	38.2	5.81	1.61	7.81*	8.4	-1.2	.9	-1.82
09C08	2009	yes	no	15.9	3.76**	1.98**	3.91**	5.83**	1	-.9	-0.36
13C08	2009	yes	no	40.8	3.34*	2.34**	3.59*	4.9*	-.7	.6	1.37
09C13	2009 <sup>γ</sup>	yes	no	49.3	17.38**	6.86	21.55**	25.19**	5.8	7.3	0.09
08C44	2009	yes	weak	51.8	0.67	1.32	0.47	0.43	1.6	1.7	0.37
01C14	2009	yes	sizeable	55.5	4.19	2.98	4.29	6*	1.3	.3	3.47
13C12	2009	yes	sizeable	66.7	6.65**	4.87**	7.14**	8.9*	2.5	1.7	2.42
06C12	2009	yes	sizeable	18.2	13.73**	5.41*	15.4**	21.97**	12.4	7	-0.09
09C07	2009	yes	sizeable	33.7	8.07*	3.47	7.54	14.14**	8.7	9	-0.19
02C06	2012	yes	no	32.6	6.03	2.06	10.55	-	8.2	2.5	-1.58
10C05	2012	yes	no	4.4	1.18	1.36	1.14	-	.5	-.3	0.4
02C11	2012	yes	no	53.9	-4.48	-4.77	-4.89	-	-7.1	-6.8	0.79
08C20	2012	yes	no	33.5	2.67	0.74	4.94	-	13.4	9.8	1.79
08C31	2012	yes	no	2	-0.21	-0.41	-0.14	-	-.1	.1	0.15
08C33	2012	yes	no	6	-2.06	-1.74	-2.52	-	.9	-.5	-0.19
02C02	2012	yes	no	13.5	3.09	1.03	5.22	-	7.3	4.9	-2.08
13C04	2012 <sup>β</sup>	yes	no	1.7	-0.44	-0.34	-0.57	-	.2	-.1	-0.24
08C14	2012	yes	no	62.2	3.46	2.84	3.69	-	.1	-.1	2.45
21C06	2012	yes	no	23.5	-0.61	-1.28	0.26	-	14	8.1	-2.33
12C05	2012	yes	no	10.8	4.25	2.49	6.05	-	-1.4	-.1	1.7
08C40	2012 <sup>β</sup>	yes	weak	27.8	-1.68	-1.01	-2.4	-	1.8	.2	-2.6
08C58	2012	yes	weak	10	2.89*	0.64	5.04**	-	0	0	0.12
07C14	2012	yes	sizeable	9.5	3.54**	1.46	5.7**	-	2.5	.2	0.68

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column signals when the policy starts. "no" (resp. "weak" / "sizeable") loose prior approval policy means that less than 10% of the stays of the DRG (resp. between 10% and 70% / more than 70%) that benefited from the loose prior approval policy. DRGs are sorted in increasing order of  $T_d$  and of this latter variable.

**Addendum:**

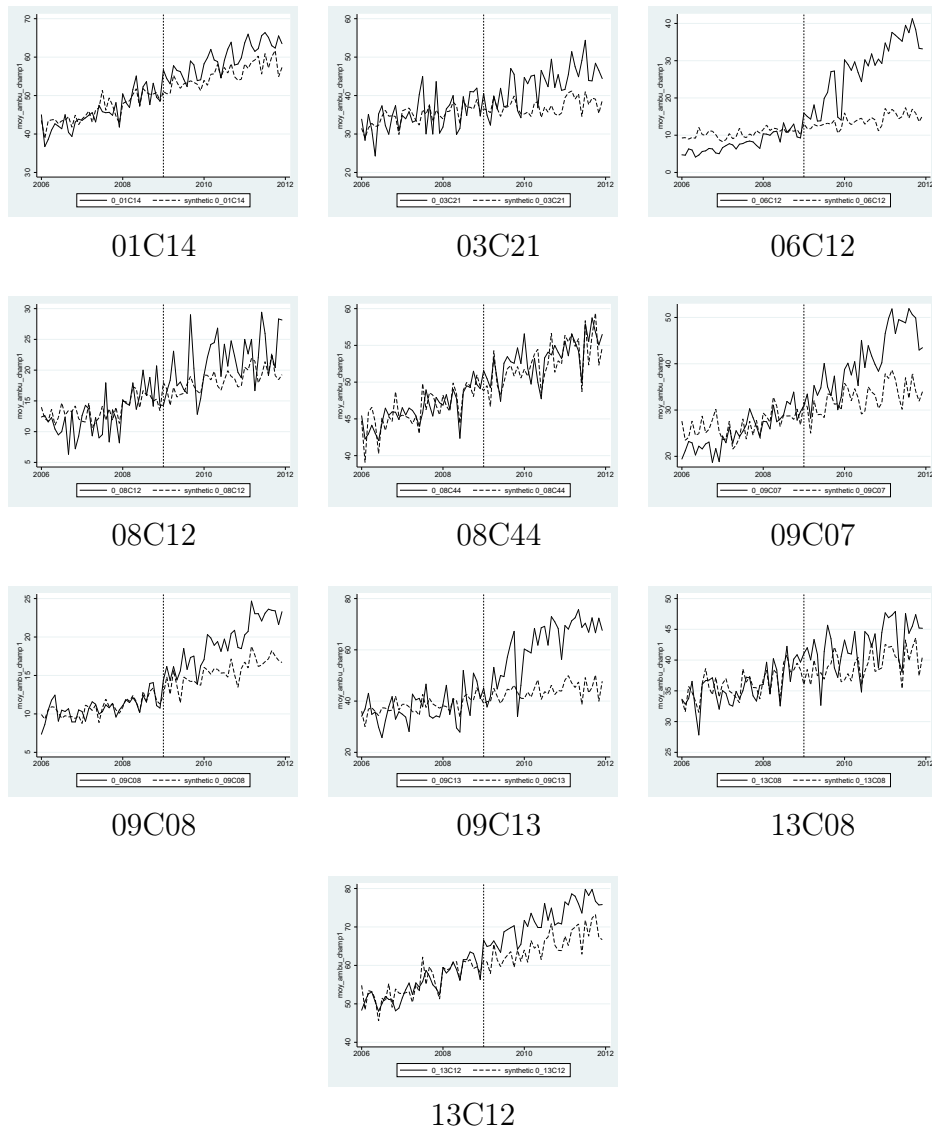
<sup>α</sup>: DRG 08C12 has no single tariff from 2010 on ; one may note that no significant effect is found for this DRG.

<sup>β</sup>: DRGs 13C04 and 08C40 have no single tariff in 2013.

<sup>γ</sup>: DRG 09C13 has a weak loose prior approval policy in 2010 and 2011 (respectively at 10% and 12%).

One may note however that the estimated effects those years are almost twice the number of stays concerned by the weak loose prior approval policy.

Figure 2: Rates of same-day surgery for DRGs with the single tariff policy starting in 2009 (public sector)



**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied.

**Source :** ATIH, PMSI-MCO, treatments by the authors.

sector. But we cannot dismiss the single tariff may have an effect also in this sector when it occurs in combination with the monitoring policy.



Table 3: Estimates in the private sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
12C08	2009	yes	no	96.2	2.62	1.95	3.81	3.03	6.5	4.5	1.22
03C15	2009	yes	no	94.8	0.77	0.4	1.69	0.69	4.5	3.5	0.71
03C21	2009	yes	no	40.2	-1.59	-0.43	-2.04	-2.98	2.5	-.4	-2.04
09C08	2009	yes	no	22.9	-0.05	1.11	-0.05	-1.1	.5	-.4	-0.47
13C17	2009	yes	no	16.8	0.93	0.87	0.6	1.29	.2	-.2	1.29
08C28	2009	yes	no	37	1.36	0.51	1.91	0.72	-.2	-3.1	-0.12
09C13	2009	yes	no	79.9	-0.45	-0.62	-1.13	0.16	.9	.4	1.8
13C08	2009 <sup>γ</sup>	yes	no	59.7	5.67*	2.47	7.22**	7.91*	-.1	.1	0.24
12C05	2009 <sup>α</sup>	yes	weak	5.8	2.28	0.1	1.68	5.03*	.2	-.1	-0.06
08C44	2009	yes	weak	83.4	2.15	1.13	2.77*	2.97	-.2	.1	0.59
13C20	2009	yes	weak	92	2.31	0.91	3.53	3.46	5.2	7.6	0.27
12C13	2009	yes	sizeable	97.4	3.31	3.04	4.46	3.46	6.5	4.5	-0.13
13C12	2009	yes	sizeable	77.9	8.18*	4.45*	9.52**	11.78*	5.8	-.2	0.75
01C14	2009	yes	sizeable	68.9	5.35*	3.54*	6.13**	7.03*	4.2	-1.5	-1.44
06C12	2009	yes	sizeable	15.9	17.73***	4.03**	20.32***	30.45***	14.1	-4.5	-2.68
09C07	2009	yes	sizeable	46.3	11.18*	5.03*	11.83**	18.29*	5.6	4.1	0.76
10C12	2012	yes	no	.3	0.34**	0.12	0.56**	-	0	0	0.04
10C10	2012	yes	no	2.8	0.36	-0.31	0.96	-	.4	-.2	0.37
02C06	2012	yes	no	76	3.48	0.53	6.74	-	.4	-1.4	3.02
10C07	2012	yes	no	3.8	-0.94	-0.76	-1.53	-	.5	-.2	-0.66
08C33	2012	yes	no	12.7	1.68	1.08	2.32	-	1.3	-1	0.79
02C11	2012	yes	no	69.9	-8.55	-6.79*	-10.33	-	.9	-2.2	-0.19
08C20	2012	yes	no	79	3.84	4.12	3.59	-	-.2	.2	-2.32
08C04	2012	yes	no	5.3	1.3	0.03	2.29	-	-2.3	-.3	0.99
13C04	2012 <sup>β</sup>	yes	no	3.2	0.12	-0.24	0.42	-	.1	-.1	-0.23
08C31	2012	yes	no	6	1.4	1.2	1.67	-	.1	.2	-0.01
21C06	2012	yes	weak	66.9	2.45	2.34	2.64	-	1.7	-1.3	-0.06
06C15	2012	yes	weak	6.2	0.14	0.14	0.28	-	.4	-.2	0.7
08C14	2012	yes	weak	75.6	1.57	0.87	2.1	-	1.1	-.9	-1.42
08C58	2012	yes	weak	8.8	2.54**	0.51	4.61**	-	.6	-.2	0.1

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column signals when the policy starts. "no" (resp. "weak" / "sizeable") loose prior approval policy means that less than 10% of the stays of the DRG (resp. between 10% and 70% / more than 70%) that benefited from the loose prior approval policy. DRGs are sorted in increasing order of  $T_d$  and of this latter variable.

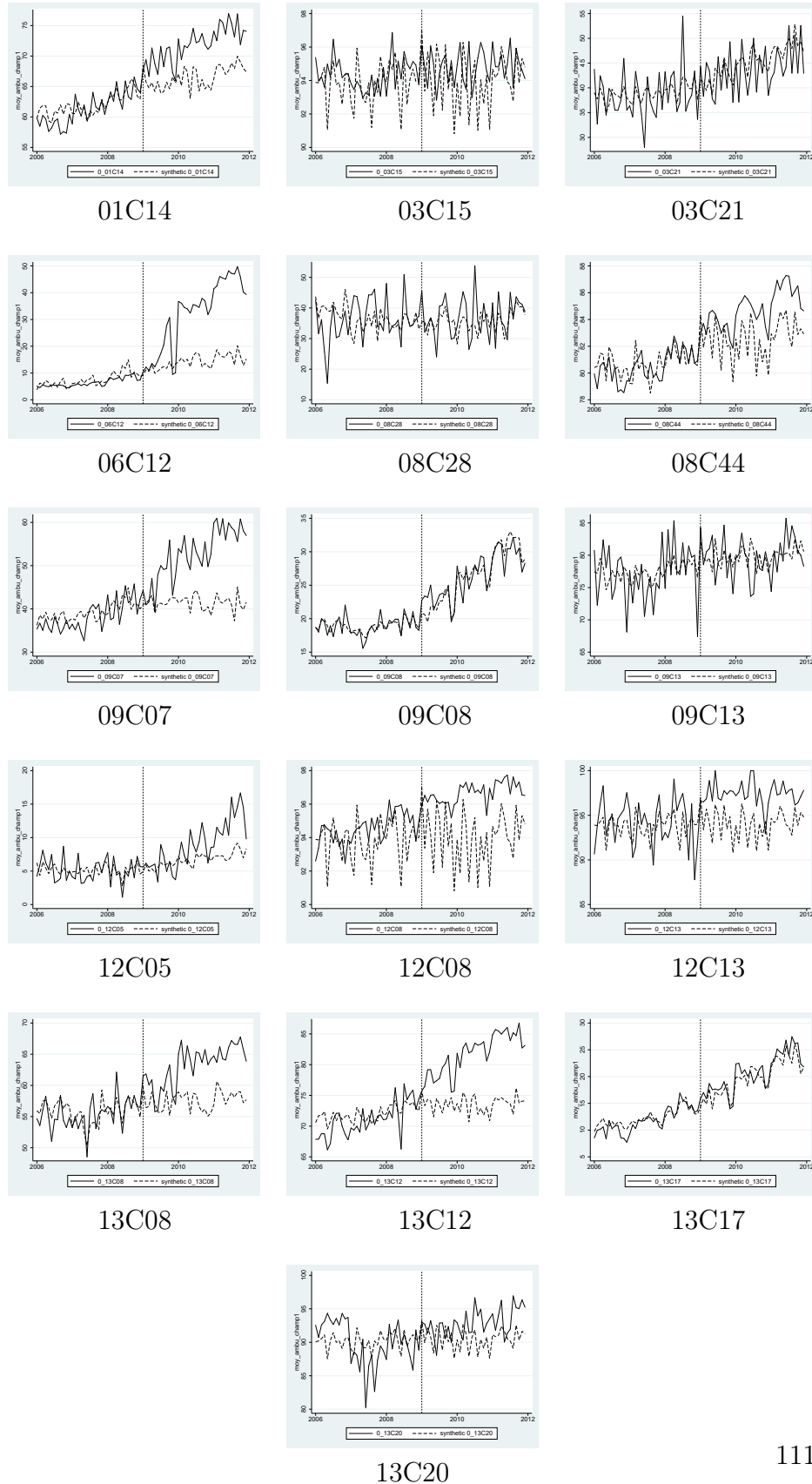
**Addendum:**

<sup>α</sup>: DRG 12C05 has no single tariff in 2010 and 2011 ;

<sup>β</sup>: DRG 13C04 has no single tariff in 2013;

<sup>γ</sup>: DRG 13C08 has a a weak loose prior approval policy in 2010 and 2011 (respectively at 11% and 11%).

Figure 3: Rates of same-day surgery for DRGs with the single tariff policy starting in 2009 (private sector)



Three general observations can be drawn from this first analysis in both the public sector and the private sector.

Firstly, when the single tariff is in place, there is no observable effect in some cases, especially in the private sector. This does not contradict the early findings on the BPTs in England, where their effects seem to have been heterogeneous. Still, it seems a necessary contribution to the literature, as, to the best of our knowledge, the only published paper in a setting really similar to ours (i.e. Allen et al. (2016)) highlights a case where the BPT policy was highly effective.

Secondly, whereas our results point to the single tariff having an effect by itself for DRGs incentivized in the public sector from 2009 onward, such evidence is scarcer in the private sector. This difference deserves further investigation. As rates of same-day surgery were higher in the private sector than in the public sector, it might have yielded less room for improvement in the private sector. Alternatively, an interaction with the monitoring policy might have been necessary in the private sector for the single tariff to have an effect by itself.

Thirdly, in both the private sector and the public sector, evidence regarding the DRGs incentivized from 2012 onward is (at best) scarce. Beyond the financial incentive that the single-tariff policy brings, another channel through which this policy may have had an observable effect is in making public hospitals focus their efforts on the development of same-day surgery for a small number of specific DRGs. This channel would have become less effective when the policy was extended in 2012 and thereafter.

## 5.2 Rates of same-day surgery for DRGs starting with only the prior approval policy

Of the 13 DRGs in the public sector for which the prior approval policy applied from 2008 or 2009 on, without any single tariff, and for which we are able to build a reasonable synthetic control, a significant effect appears in 7 cases (see Table 4).<sup>38</sup> Discarding 3 DRGs<sup>39</sup> for which the effect starts being significant only from the year when a single tariff starts being applied, there are 4 DRGs with a positive and significant effect<sup>40</sup> out of 10 DRGs.

For the 4 DRGs in the public sector without any single tariff, but for which the prior approval policy applied from 2012 on, no significant effect appears.

Overall, this provides evidence that in the public sector, the prior approval policy alone has occasionally significantly increased same-day surgery rates of targeted DRGs. This is in itself a significant result as we studied a loose prior approval policy. And indeed, the first year effects that we find are much higher than the corresponding percentages of stays under effective prior approval (which was barely implemented in the first year of its application). This is *ex-post* evidence of a signal effect. Hospitals did not wait to be subject to prior approval for a list of acts in the prior approval policy to consider raising their corresponding rates of same-day surgery. The absence of effect for those incentivized starting in 2012 may suggest that this signal effect vanished in further extensions of the policy.

<sup>38</sup>The corresponding graphs for those incentivized from 2008 or 2009 on are in Figure 4, and in Figure 8 of Annex B for those incentivized starting in 2012.

<sup>39</sup>There are: DRG 08C38: *Other knee arthroscopy*, DRG 02C05: *Interventions on the lens with or without vitrectomy.*, DRG 08C45: *Arthroscopic meniscectomy.*

<sup>40</sup>These 4 DRGs are DRG 06C10: *Restorative procedures for hernias and eventrations, age under 18.*, DRG 08C60: *Interventions on the wrist other than tenosynovectomies*, DRG 06C14: *Interventions on the rectum and anus other than rectal resections.*, DRG 13C10: *Tubal ligation by laparoscopy or laparoscopy.*

Table 4: Estimates in the public sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo effect in $T_d - 1$
08C54	2008	no	weak	9.4	2.06	1.39	2.83	2.01	.9	-.3	-1.63
03C17	2008	no	weak	42.7	-0.94	-1.23	-0.96	-1.07	3.6	3.2	-5.32
08C38	2008 <sup>δ</sup>	no	sizeable	44.7	9.58	5.78	10.5*	13.54*	6.4	3.9	-0.98
02C05	2008 <sup>δ</sup>	no	sizeable	65.3	13.93*	7.75	18.13**	17.84*	8.5	7.6	2.36
08C45	2008 <sup>δ</sup>	no	sizeable	60.1	12.37**	5.97	14.88***	17.61**	10.4	11.5	1.79
06C10	2009 <sup>ε</sup>	no	weak	52.2	5.5**	4.91**	7.08**	5.67	-.9	1.1	1.19
08C60	2009	no	weak	29	2.9	1.92	2.4	5.11*	1.7	-1	-0.23
02C08	2009	no	weak	61.7	1.8	2.3	2.38	1.4	.7	1.5	0.51
08C46	2009	no	weak	54.1	0.98	1.06	1.25	1.05	.3	1.3	1.1
06C14	2009	no	weak	12.7	1.79**	0.67**	1.64**	3.26**	.8	-.8	-0.46
13C19	2009	no	weak	11.7	1.55	-0.45	0.65	4.49	-.2	4.9	0.36
13C11	2009	no	sizeable	64.8	5.29	3.03	4.3	8.81	2.8	1.8	1.53
13C10	2009	no	sizeable	25	11.45**	4.66	11.36**	19.44**	3.4	4.2	-0.28
13C05	2012	no	weak	1.5	0.71	0.57	0.85	-	-.1	.1	0.16
13C06	2012	no	weak	24.4	-3.29	-3.36	-3.45	-	10	27	-1.98
11C07	2012	no	weak	25.8	0.41	-0.72	1.2	-	2.6	-.9	-0.47
03C06	2012	no	weak	15.2	-2.63	-4.02	-1.76	-	9.4	7.8	-5.76

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

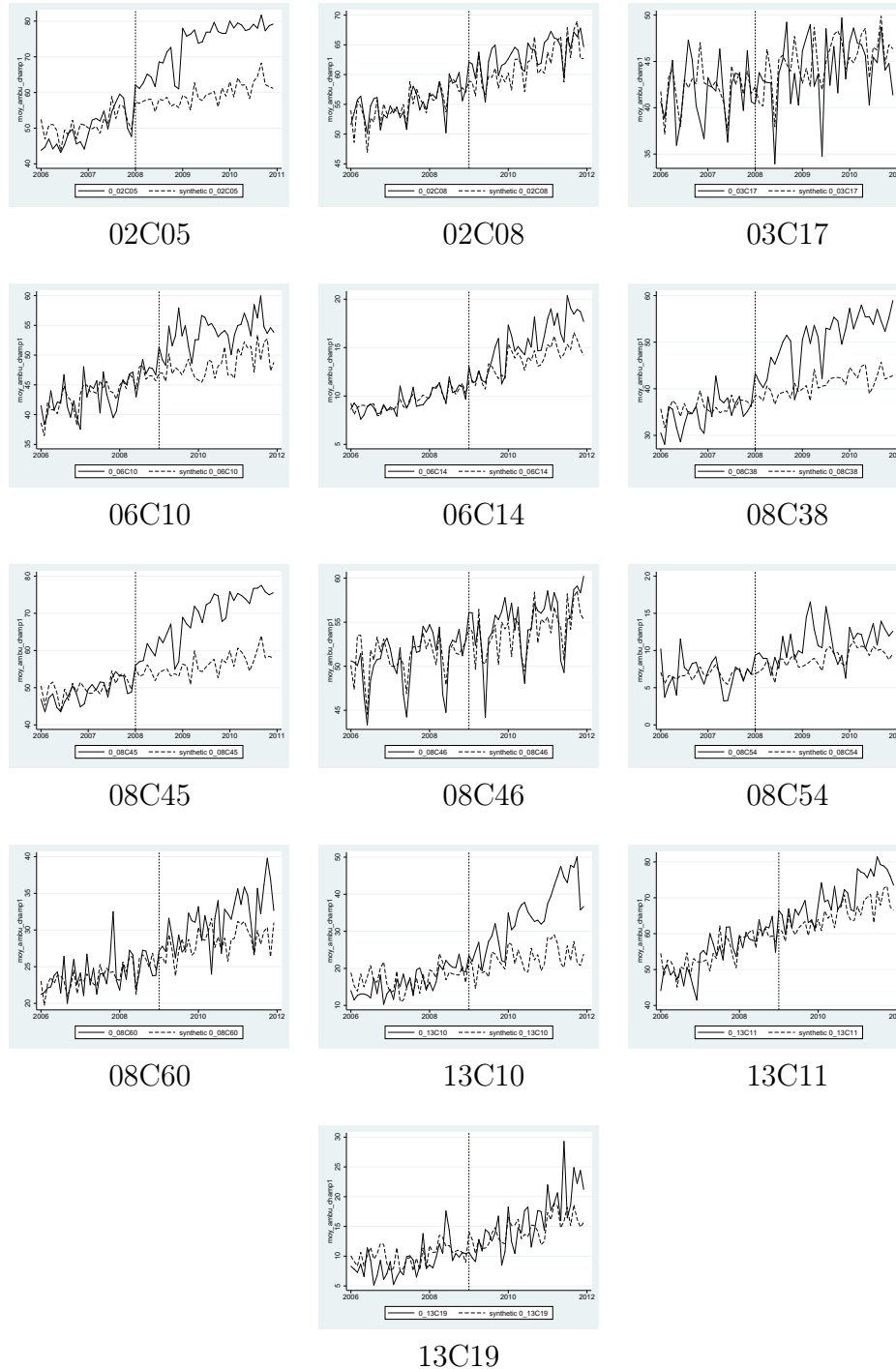
Column signals when the policy starts. "no" (resp. "weak" / "sizeable") loose prior approval policy means that less than 10% of the stays of the DRG (resp. between 10% and 70% / more than 70%) that benefited from the loose prior approval policy. DRGs are sorted in increasing order of  $T_d$  and of this latter variable.

**Addendum:**

<sup>δ</sup>: DRGs 02C05, 08C38 and 08C45 have had a single tariff from 2009 on;

<sup>ε</sup>: DRG 06C10 has had a single tariff from 2010 on. One may note yet that the significant effect appears as soon as 2009.

Figure 4: Rates of same-day surgery for DRGs with only the prior approval policy in 2008 or in 2009 (public sector)



**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line.

The vertical line in March 2009 stands for the date when the policy applied.

**Source :** ATIH, PMSI-MCO, treatments by the authors.

Of the 14 DRGs in the private sector without any single tariff, but for which the prior approval policy applied from 2008 or 2009 on, and for which we are able to build a reasonable synthetic control, a positive and significant effect appears in 4 cases<sup>41</sup> (see Table 5).<sup>42</sup> But one of these cases should be discarded from the analysis as the effect appears in fact only in 2010, as the DRG starts benefiting from the single tariff as well.

No effect is found among the 6 DRGs incentivized from 2012 on.

Overall, the monitoring policy alone may have had a strong effect, notably in the public sector, but in some cases only.

---

<sup>41</sup>This concerns DRG 03C17: *Interventions on the mouth*, DRG 06C10: *Restorative procedures for hernias and eventrations, age under 18*, DRG 02C13: *Interventions on oculomotor muscles, age <18 years.*, and DRG 13C19: *Interventions for infertility or reproductive care reasons.*

<sup>42</sup>The corresponding graphs for those incentivized in 2008 or in 2009 are in Figure 5, and in Figure 9 of Annex B for those incentivized from 2012 on.

Table 5: Estimates in the private sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
08C34	2008	no	weak	.3	-0.18*	-0.06	-0.05	-0.44**	0	0	-0.02
03C06	2008	no	weak	38.4	0.23	1.61	-0.62	-0.15	-.3	.8	4.98
08C54	2008	no	weak	16.1	1.26	0.92	1.56	1.63	.5	-.5	1.54
03C17	2008	no	weak	70.9	4.7	0.75	5.99*	7.57	2.7	-1	-2.15
06C19	2009	no	weak	9.7	-0.56	-0.3	-0.36	-0.98	.8	-.6	-1.26
06C14	2009	no	weak	23.3	-0.95	-0.17	-1.1	-1.66	1	-1	0.08
08C59	2009	no	weak	84.4	1.39	0.56	2.4	2.03	3.9	-.4	0.5
21C04	2009	no	weak	72	2.5	1.39	4.06	2.8	-.2	.2	-0.62
02C08	2009	no	weak	77.3	-0.15	-0.77	-0.1	0.76	2.5	-.9	-2.35
08C46	2009	no	weak	76	1.03	0.09	1.64	1.84	-.2	.3	1.29
06C10	2009 <sup>ε</sup>	no	weak	55.4	7.22*	-0.26	9.32*	12.63*	2.2	-1.1	2.48
02C13	2009	no	weak	36.5	6.46	1.04	5.18	13.28*	11.8	-4.5	-0.95
13C19	2009	no	weak	7.1	3.25*	-0.35	2.89	7.36*	.1	.5	-0.51
08C60	2009	no	weak	53.5	-0.31	-0.4	0.29	-0.63	1.4	-1	-0.28
13C05	2012	no	weak	1.3	-0.29	-0.51	-0.2	-	0	0	-0.25
09C14	2012	no	weak	18.6	3.22	1.72	5.39	-	.5	0	-1.48
08C35	2012	no	weak	5.9	0.02	-0.23	0.22	-	-.1	0	0.41
08C43	2012	no	weak	61.2	-0.05	-0.98	0.7	-	.5	-.4	0.16
03C07	2012	no	weak	27.7	1.42	0.95	2.01	-	4	-.1	-0.92
11C07	2012	no	weak	54.6	0.08	0.97	-0.37	-	.4	-2.7	-1.46

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

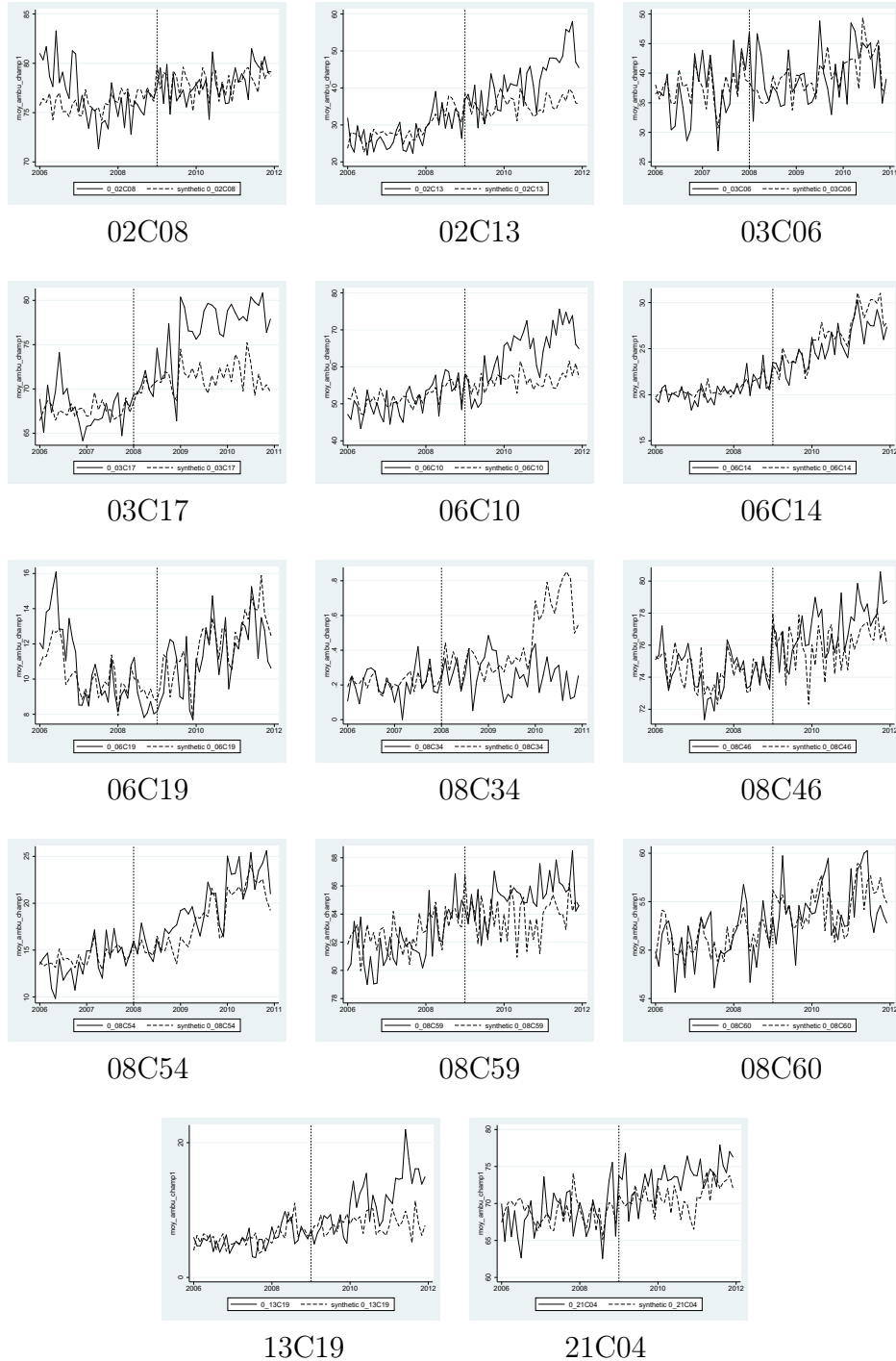
Column signals when the policy starts. "no" (resp. "weak" / "sizeable") loose prior approval policy means that less than 10% of the stays of the DRG (resp. between 10% and 70% / more than 70%) that benefited from the loose prior approval policy. DRGs are sorted in increasing order of  $T_d$  and of this latter variable.

**Addendum:**

<sup>ε</sup>: DRG 06C10 has had a single tariff from 2010 on.



Figure 5: Rates of same-day surgery for DRGs with only the prior approval policy in 2008 or in 2009 (private sector)



**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied. **Source :** ATIH, PMSI-MCO, treatments by the authors.

### 5.3 Magnitude of the effects

In sections 5.1 and 5.2, we have focused on the question of whether the single tariff policy or the prior approval policy triggered increases in same-day rates. We now turn to the question of the magnitude of the estimated effects, and of their representativeness.

We focus first on the DRGs in the public sector that benefited from either or both policies from 2008/2009 onward. There are 45 of them, that account in 2011 for 710,000 stays out of the 2.2 million stays of the field of interest in the public sector (that is 32% of the field of interest in the public sector, where the field of interest includes all surgical stays save surgical operations on pathological pregnancies and births). However we have to dismiss some of them due to lack of information regarding same-day prospects, or because we could not build a reasonable synthetic control. We are eventually able to study 23 DRGs, that stand for 470,000 stays in 2011 - that is 66% of the field of incentivized DRGs, and 21% of the field of interest in the public sector.

Over these 23 DRGs, we find a significant increase in the same-day rate for 15 of them, that account for 360,000 stays in 2011. Applying the cumulative increase after three years of any policy to the number of stays in 2011, we assess that the policies would have been responsible for 54,000 same-day stays in 2011. This is equal to 11.5% of stays of the 23 incentivized DRGs that we could study, or 2.4% of the field of interest in the public sector.<sup>43</sup>

Applying this reasoning to the DRGs that benefited from either or both poli-

---

<sup>43</sup>In the private sector, the significant increases on the 30 DRGs incentivized from 2008/2009 on and that we could study are associated with a 3-year increase of 4.7% on their field, that is 0.9% of the field of interest in the private sector.

cies from 2012 onward, we estimate that the policies would have caused a 2-year increase of 0.7% of stays of the 18 incentivized DRGs we could study, or 0.1% of the field of interest in the public sector.<sup>44</sup>

These estimates seem somewhat below governmental expectations. In their optimistic scenario, Bert et al. (2014) assessed that the same-day rate could increase in the public sector from 31.3% in 2013 to 56.2% in 2018. This would require 3-year increases in same-day rates of 15% over the total field of interest in the public sector. This is not in line with our above estimates, even if the policies were getting close to a 11.5% 3-year increases over the whole field (as observed on the field of the studied DRGs with an incentive from 2009 onward).

Moreover, even this 11.5% figure seems overestimated and is assuredly very imprecise: Although we have not computed a confidence interval for each estimate, a crude measure of the possible imprecision can be obtained by subtracting the difference in (low) mid-term prospects of same-day rates from the 3-year effect we get for each incentivized DRG. Proceeding as before, we can re-assess a lower threshold for the 3-year increase over the overall field of all incentivized DRGs that we can study. This enables us to take into account the residual difference in potential achievable levels of the incentivized DRGs and their synthetic controls. As the 11.5% increase shrinks to 6.2% with this correction, it may be concluded that this initial figure is probably overestimated.

Another area that may further add to the lack of precision is that such an assessment is very sensitive to few DRGs. For instance, applying the above correction to only the three DRGs incentivized from 2008 onwards and for which we find strong

---

<sup>44</sup>In the private sector, the significant increases on the 20 DRGs incentivized from 2012 onwards and that we could study are associated with a 2-year increase of 0.4% on their field, that is 0.07% of the field of interest in the private sector.

effects, we already observe that the 11.5% increase shrinks to 7.9%. This is partly because the above correction is large for these three DRGs, and also because they account for very large numbers of stays - 197,000 stays in 2011 (that is 42% of the stays of the incentivized DRGs we could study).

Conversely the non-significant increases obtained for the DRGs incentivized from 2012 onwards could be seriously underestimated. As the policies had been expanded in 2012 to many new DRGs, stakeholders could have anticipated further expansion of these policies and even a generalisation of the single tariff: they might have expanded their effort beyond the incentivized DRGs<sup>45</sup>. These potential "anticipatory effects" - that are also well-known to the practitioners of Randomized Control Trials (see for instance Heard, O'Toole, Naimpally, and Bressler (2017) ) - would downward bias the results as they would make the studied policies affect the control group, and hence also the synthetic controls.

These hypotheses are closely related to the question of the representativeness of the results we obtained for the DRGs incentivized in 2008/2009. We present below the average same-day rate (as of 2006/2007)<sup>46</sup> and the average lag of their same-day rate at the year they received an incentive behind the (low- and high-) same-day prospects, grouping DRGs by sector and by year of their first incentive. One observes a (roughly) decreasing pattern of the lag behind the same-day prospect at the first year of the incentive. This suggests that DRGs which were the first to be targeted by the policies were furthest away from their estimated potential of same-day surgery rate. Since, the estimated effects among the DRGs incentivized in 2008/2009 increase with such lag (as displayed in Annex G), one

---

<sup>45</sup>Very similarly, their effort for the incentivized DRGs could have transverse effects benefiting also to the DRGs that are not incentivized ; this is all the more likely if there are large sunk costs to develop and maintain same-day units.

<sup>46</sup>This is an unweighted average.

might expect that the effects of either policy decrease on average as it is applied to DRGs with less potential.

Table 6: **Representativeness of the DRGs incentivized in 2008/2009**

first year of the incentive	public sector				private sector			
	Number of DRGs	same-day rate in 2006/2007	lag behind the low prospect	lag behind the high prospect	Number of DRGs	same-day rate in 2006/2007	lag behind the low prospect	lag behind the high prospect
2008	8	23%	20%	25%	8	37 %	12%	19%
2009	29	41%	19%	27%	36	50 %	9%	17%
2012	32	10%	12%	19%	32	17 %	0.8%	7.5%
2013	16	22%	10%	14%	21	31 %	-7%	-2%
2014	62	4%	1%	5%	52	9 %	-4%	-1%

**Lecture:** The lag behind the low prospect (resp. the high prospect) is the difference between the low mid-term prospect of same-day rate (resp. the high mid-term prospect of same-day rate), and the same-day rate at the first year of the incentive.

**Note:** The mid-term prospect of same-day rate we have does not distinguish between the public and the private sector, which explains that the average "lag" can be negative in the private sector.

**Observations:** Surgical DRGs for which we have a mid-term same-day prospect.

## 5.4 Impacts on other variables

We study the evolution of certain complementary indicators such as re-hospitalisation and the total number of stays.

To estimate an effect on these complementary variables, we proceed in the same way as for the same-day surgery rate : a new synthetic control is re-estimated (but we still use the two same-day prospects as control variables).

In tables 7 to 9 we report the results of the estimations made on the complementary variables of interest. To work on comparable cases, we focus on the DRGs in the public sector for which a significant and positive effect of the incentive has been found.<sup>47</sup> We do not present the estimates that did not pass the conditions outlined in section 4.5.

The estimations regarding the percentage of stays of severity 1 (which are expected to be the stays replaced by same-day surgery due to each measure) in most cases give opposite results to those for the same-day surgery rate and are of similar magnitude. The effects associated with the stays of severity 2 are almost all negative, and none is positive and significant, thereby dismissing any worry of upcoding to severity level 2 due to the measure (one might have had such concerns as the tariff for stays of severity 1 probably decreased relative to the tariff for stays of severity 2).

A somewhat similar picture emerges regarding 2-night stays and 3-night stays (whose associated coefficients are almost all negative), but not regarding 1-night

---

<sup>47</sup>For the DRGs incentivized in 2009, we added the condition that this effect had to be persistent.

stays, whose coefficients are sometimes positive and sometimes negative. Various signs could be possible in 1-night stay coefficients due to the measures if they encouraged surgeons to plan more same-day surgery (notably instead of 1-night stays), and if some of this additional planned same-day surgery eventually required one night at the hospital (or additional care at the hospital on the following day due to complications). The total effect on 1-night stays would thus be ambiguous and would depend on the DRG.

The measures do not seem to have led to an increase in the total number of stays. There is one single significant and positive effect, but two years later, the corresponding coefficient is not significant and negative.<sup>48</sup> Hence, there is no pattern showing a general significant increase in the total number of stays.

Regarding re-hospitalisation, the associated coefficients are sometimes negative, and sometimes positive. There are negative and significant effects for DRG 06C14 but the corresponding placebo is also negative and significant, so it should not be considered. There is only one positive and significant coefficient for the 30-day re-hospitalisation rate of DRG 02C05, that is *Interventions on the lens with or without vitrectomy*, applying to cataracts. Yet, as noted by Yilmaz and Vuagnat (2015b), cataract treatment usually requires two operations, one for each eye. And with same-day surgery, the second operation can be performed quicker - within 2 weeks. In this case, an increase in 30-day re-hospitalisation should not be considered as a proxy for downgrading quality.

---

<sup>48</sup>This concerns DRG 06C14: *Interventions on the rectum and anus other than rectal resections*.

Table 7: Estimates on the complementary variables for the incentivized DRGs with a robust effect on the same-day surgery rate

	name	09C08	13C08	09C13	13C12	06C12	09C07	06C10	08C60	06C14	13C10	08C38	02C05	08C45
	$T_d$	2009	2009	2009	2009	2009	2009	2009	2009	2009	2009	2008	2008	2008
	sector	public	public	public	public	public	public	public	public	public	public	public	public	public
	monitoring	no	no	no	large	large	large	weak	weak	weak	large	large	large	large
	single tariff	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no	no
% day-care	Total effect	3.8**	3.3*	17.4**	6.7**	13.7**	8.1*	5.5**	2.9	1.8**	11.4**	9.6	13.9*	12.4**
	1-year effect	2**	2.3**	6.9	4.9**	5.4*	3.5	4.9**	1.9	0.7**	4.7	5.8	7.7	6
	2-year effect	3.9**	3.6*	21.5**	7.1**	15.4**	7.5	7.1**	2.4	1.6**	11.4**	10.5*	18.1**	14.9***
	3-year effect	5.8**	4.9*	25.2**	8.9*	22**	14.1**	5.7	5.1*	3.3**	19.4**	13.5*	17.8*	17.6**
	diff in low b.	1	-7	5.8	2.5	12.4	8.7	-9	1.7	.8	3.4	6.4	8.5	10.4
	diff in high b.	-9	.6	7.3	1.7	7	9	1.1	-1	-8	4.2	3.9	7.6	11.5
	Plac. eff.	-0.4	1.4	0.1	2.4	-0.1	-0.2	1.2	-0.2	-0.5	-0.3	-1	2.4	1.8
	value in $T_d$	15.9	40.8	49.3	66.7	18.2	33.7	52.2	29	12.7	25	44.7	65.3	60.1
% 1-level	Total effect	-3.3*	-1	-15.5	-8.1	-11.8***	-6.6	-2.3	-0.9	-0.9	-12.1*	-8.6	-	-13.3*
	1-year effect	-1.8	-1	-6.1	-5.8	-5.3**	-2.9	-2.5	-1.1	0.1	-8.4*	-4.6	-	-7*
	2-year effect	-3.5*	-1.2	-19.1*	-8.8	-12.8***	-5.6	-2.9	0.4	-0.9	-12.8*	-9.8	-	-15.7*
	3-year effect	-5.1**	-1.4	-22.8	-10.6	-18.5***	-12.3	-2.3	-2.4	-2.3	-16.2*	-12.5	-	-18.5*
	diff in low b.	.8	-1	6.8	3.6	2.9	9.7	-3	1.3	.3	-5.3	6.3	10.1	11.4
	diff in high b.	-7	.7	5.7	2.6	.8	10.3	.6	-8	-3	7.9	2.9	7.2	11.4
	Plac. eff.	0.9	-2.3	0.4	-3	0.3	0.1	-1.4	1.6	0.7	-3.3	0.5	-	-2.5
	value in $T_d$	81.8	55.2	49.5	32.5	67.8	65	47.1	66.8	77.9	42.4	52.4	33	38.5
% 2-level	Total effect	-	-0.7	-	-0.2	-2.8**	-	-0.1	-0.1	-1.6*	-	-	-	-
	1-year effect	-	-0.5	-	-0.2	-1.1	-	-0.2	0	-1.2**	-	-	-	-
	2-year effect	-	-0.2	-	-0.2	-3**	-	-0.2	-0.5	-1.5*	-	-	-	-
	3-year effect	-	-1.2	-	-0.2	-4.6**	-	0.1	0	-2.1*	-	-	-	-
	diff in low b.	0	0	.1	0	0	-1.3	-1	.7	.1	-3.2	.1	.6	.5
	diff in high b.	0	0	.1	.2	1.6	3.7	.3	-9	-2	4.5	.3	2.6	7.5
	Plac. eff.	-	-0.1	-	-0.1	-0.3	-	-0.2	0.1	-1.3***	-	-	-	-
	value in $T_d$	1.6	2.8	.9	.7	10.7	1.1	.5	3	6.6	.4	1.8	.3	.4



Table 8: Estimates on the complementary variables for the incentivized DRGs with a robust effect on the same-day surgery rate

	name	09C08	13C08	09C13	13C12	06C12	09C07	06C10	08C60	06C14	13C10	08C38	02C05	08C45
%	1-night													
	Total effect	7.6	2	0.3	-3.9	1.9	0.4	-2.3	-0.9	4.2*	3	-3.8*	-	-6.9**
	1-year effect	7.1	1.3*	2.5	-1.6	1.9	0.2	-1.6	-0.4	1.6	3.8	-1.7	-	-3
	2-year effect	6.4	2.1	0.6	-4.7	1.8	0.3	-2.3	-0.8	4.5**	4.1	-4.6*	-	-8**
	3-year effect	9.9	2.6	-1.2	-5.9	2.4	0.5	-3.1	-1.6	6.9*	1.1	-5.4*	-	-10.3**
	diff in low b.	-15.9	-7	2.4	3.6	.4	-2.6	.6	1.8	1.3	-3	1.4	4.3	9.6
	diff in high b.	-14	2.5	3.2	5.6	1.6	8.4	.1	-1.7	-2.1	16.8	4.8	6.3	16.2
	Plac. eff.	6.9*	0.6	-1.5	-0.5	0.9	0.5	-0.7	0.3	1.3*	3.5	-0.9	-	-0.7
value in $T_d$	40.2	31.2	19.7	19.5	14.2	18.8	19.1	20.4	24.9	32.1	20	17.4	21.1	
%	2-night													
	Total effect	-2.7*	-2.5	-	-3.4	-1.2	-	-1.3	0.7	-1.3	-	-4.1	-	-3.9
	1-year effect	-2	-2	-	-3.6*	-0.1	-	-2.5	0.4	-1	-	-1.9	-	-1.7
	2-year effect	-2.2	-2.3	-	-3	-1	-	-1.3	0.8	-1.5	-	-4.8	-	-4.5
	3-year effect	-4.1*	-3.3	-	-4.2	-2.8	-	-0.7	1.1	-1.6	-	-6.1	-	-6
	diff in low b.	.6	-2	26.2	1.1	1.1	24.9	1.6	1.5	.7	16.6	5.8	14.2	7.4
	diff in high b.	-7	.7	19.8	2.8	2	34.2	2.4	-7	-6	32.6	6.2	15.1	13.9
	Plac. eff.	-0.6	-1.6*	-	-1.9	2.2	-	-1.4	3	0.9	-	2.8	-	1.3
value in $T_d$	26.2	16.5	25	10.7	27.5	35.1	22.6	22.4	26	34.5	19.5	14.2	14.4	
%	3-night													
	Total effect	-2.7	-0.8	-	-	-	-	-	0.5	-2*	-	-1	-	-
	1-year effect	-1.5	-0.6	-	-	-	-	-	0	-0.3	-	-1.1	-	-
	2-year effect	-2.6	-0.9	-	-	-	-	-	1.4	-2.3*	-	-0.6	-	-
	3-year effect	-4.4*	-0.9	-	-	-	-	-	0.1	-3.5*	-	-1.3	-	-
	diff in low b.	.2	-1.3	1.6	-8	19.2	-1.3	-1.1	2.7	.6	-6.2	1.9	.5	2.7
	diff in high b.	-2	1.6	1.5	1.2	2.3	6.1	1.5	-1.2	-6	9.3	1.3	2.5	7.6
	Plac. eff.	-1.6	-1	-	-	-	-	-	0.8	-0.8	-	0.1	-	-
value in $T_d$	9.7	5	3.9	1.6	19.6	8.2	3.5	13.9	15	7.1	5.2	.9	2	

Table 9: Estimates on the complementary variables for the incentivized DRGs with a robust effect on the same-day surgery rate

	name	09C08	13C08	09C13	13C12	06C12	09C07	06C10	08C60	06C14	13C10	08C38	02C05	08C45
nbr of stays relative to 2006	Total effect	0.1	0	0.3	0	-0.1	-0.2	0	0	0	-0.2	-0.1	0	-0.1
	1-year effect	0	0	0.1	0	0	-0.2	0	0	0.01*	-0.2	-0.1	0	0
	2-year effect	0.1	0	0.4	0	-0.1	-0.2	0	0	0.01	-0.2	-0.1	0.1	-0.1
	3-year effect	0.1	-0.1	0.4	-0.1	-0.1	-0.3	0	0	-0.02	-0.2	-0.1	0.1	-0.1
	diff in low b.	.1	.3	3.8	7.2	2.5	14.1	2	4.9	1.5	6	8.5	20.8	11.9
	diff in high b.	2.8	2.4	4.9	9.1	4.5	9	2.8	-1.9	-1.3	4.1	3.1	21.5	18.6
	Plac. eff. value in $T_d$	0.1 1.3	0 1.1	0 1.3	0 1.1	0 1.1	-0.1 .7	0 .9	0 1.2	0 1.1	-0.1 .8	0 .9	0 1.1	0 1
% 7-day rehospitalization	Total effect	0.1	0	-	-	0.2	-	0.1	-	-0.2	-	-	0.5	-
	1-year effect	-0.1	0.1	-	-	-0.1	-	0	-	0	-	-	0.3	-
	2-year effect	0.4	0.1	-	-	0.3	-	0.2	-	-0.2	-	-	0.3	-
	3-year effect	0.1	-0.1	-	-	0.3	-	0.2	-	-0.4	-	-	1.1	-
	diff in low b.	.1	-1	4.1	1.3	-2	.4	-3	1	.5	-5	0	7.3	4.8
	diff in high b.	.1	.7	3.9	3.3	.3	12.1	1.4	-7	-4	8.6	3.1	9.4	11.6
	Plac. eff. value in $T_d$	0 1.6	0.3 2.2	- 1.5	- 1.5	0.2 1.5	- .6	0 1.5	- 1.2	-0.8** 3	- .8	- 1.3	-0.2 3.2	- 1
% 30-day rehospitalization	Total effect	-0.2	-0.4	-	-	0	-	-0.1	-0.1	-0.8*	-	-1.1	3.8	-
	1-year effect	-0.3	-0.2	-	-	-0.2	-	-0.5	-0.4	-0.6**	-	-1.3	2.1	-
	2-year effect	0	-0.2	-	-	0.1	-	0.1	-0.2	-1.1*	-	-1.2	3.5	-
	3-year effect	-0.2	-0.7	-	-	0.1	-	0.1	0.3	-0.8	-	-0.8	6.2**	-
	diff in low b.	.2	0	1.4	.6	-2	-8	-3	.8	.3	-1.3	.4	13.2	2.4
	diff in high b.	-5	.1	2.8	.9	.2	3.7	.7	-1	-4	6	1.5	15.4	8.6
	Plac. eff. value in $T_d$	-0.6 4.3	-0.6 6.8	- 5.4	- 5	0.1 4.1	- 3.1	-0.3 3.6	0.2 3.5	-1.1*** 9.4	- 2	-0.4 4	1.4 14	- 2.7

## 6 Conclusion

We assessed the effects of two French policies aiming to encourage hospitals to steer their medical practices towards increased same-day surgery rates. To take into account the selection of DRGs that benefited from either policy, we built synthetic controls and used prospects of same-day surgery rates provided by medical learned societies. Our main results are the following.

We first show that the single tariff policy alone generally did not have a significant impact on the same-day surgery rate of the studied DRGs in the private sector. This policy occasionally had a significant impact, when the single tariff policy was accompanied by the prior approval policy. However, the policy had a significant impact in the public sector for a majority of DRGs incentivized in 2008 or 2009, irrespective of the level of the prior approval policy. In this respect, our evidence complements Allen et al. (2016)'s conclusions: "*Payers can act proactively in their price setting and might*" sometimes "*expect a substantial response from hospitals*".<sup>49</sup> At a somewhat more micro-economic level, this seems to indicate that hospital managers in the public sector can make hospital surgeons change their medical practices under some circumstances.

As a precautionary note, the single tariff policy might also have had transverse impacts that we cannot investigate and that might have caused us to underestimate the effects of the policy. Namely, the hospitals might have reacted to the general trajectory of revaluation of same-day surgery, rather than to the key moments of this revaluation, and might have reacted on same-day surgery generally and not specifically on the first DRGs explicitly concerned by the revaluation. This could explain why we rarely find significant effects neither in the public sector nor in the

---

<sup>49</sup>We also show, as done before by Verzulli et al. (2016), that the finding by Januleviciute et al. (2016) that surgical DRGs do not respond to price is context-dependent.

private sector from 2012 on. Transverse effects are all the more likely considering that the development of hospital-wide same-day surgery can generate initial costs (reorganisation, potentially real estate if there was no same-day surgery unit, etc.) that may be both significant and transverse (the same-day surgery units are probably not dedicated to the practice of the few DRGs supported by the single tariff policy).

Secondly, the same-day surgery rate of some DRGs that benefited from the (loose) prior approval policy from 2009 onwards increased significantly in the public sector. This suggests that, in some cases, the prior approval policy might have acted as a strong signal that efforts should be devoted to the listed acts, even if the hospital was not subject to this strict monitoring procedure. Such an effect could come either from learning that the same-day procedure was so reliable that it had been included in the list of targeted acts, or out of concern of being monitored. Moreover, as the burden of the monitoring procedure would have weighted on both hospital management and doctors, the policy provided incentives to the key stakeholders who could develop the practice of same-day surgery in the hospital.

Thirdly, among the groups whose same-day surgery rate increased as a result of either policy, the measure does not appear either to have led to a general rise in the total number of stays or to have increased re-admissions. There is thus no evidence that the quality of care provided decreased

As previously noted, the public sector seems more responsive than the private sector to the two studied policies, which might appear paradoxical for the single tariff policy. A first hypothesis could be that there was more room for improvement regarding the practice of same-day surgery within the public sector (we present

suggestive evidence of this in Annex G), and that the larger the opportunity for improvement, the higher the expected effect. For instance, within both sectors, the estimated effects indeed increase with the distance to the mid-term prospect, which is a measure of the potential for improvement.<sup>50</sup> A second hypothesis, based on the observation that the private sector responded only when the two policies were combined, would be that it is particularly important in this sector that incentives should target both hospital management and doctors.

Such observations call for great attention to context when assessing the efficiency of a policy. While both policies brought incentives to boost the practice of same-day surgery, maybe, for various DRGs in the private sector, there were at that time fewer opportunities to further improve same-day surgery rate. Moreover, given the lack of significant effects from both policies when they were introduced for new DRGs in 2012, incentives might not always be enough. Incentives might also benefit from complementary policies such as information campaigns like those organised locally when the cited Best Practice Tariff was launched (see Allen et al. (2016) and ANAP and HAS (2014)).<sup>51</sup> What is more, the two policies have the theoretical downside of the prospective payment system in which they take place: they focus on productivity, which could be obtained at the expense of quality. Though we did not find such a pattern in the few cases that we were able to investigate, this issue is currently being targeted by the Health Ministry, which is, for instance, assessing the re-hospitalisation rate of hospitals and comparing it with a statistical expected rate of re-hospitalisation (DGOS & ATIH, 2017).

The single tariff policy was also designed as a budget-neutral policy, the first

---

<sup>50</sup>See also Annex G.

<sup>51</sup>Note that this idea that practical advice should be delivered as well was first implemented from 2013 on, with notably the publication of reports and of a software to help hospital managers to foster same-day surgery (HAS & ANAP, 2013, 2015).

single tariffs being approximately chosen in this perspective. This decision was beneficial to all the hospitals that were already practicing a significant amount of same-day surgery, and was detrimental to the others, making it potentially even more difficult for the latter to find the financial resources to bring about changes. Ultimately, the policy may have had budget impacts if these establishments had to be bailed out. One avenue for future research could therefore be to investigate the microeconomic impacts of these kinds of policies at the hospital level.

Another avenue for improvement relates to the quantification of the uncertainty around our estimates. As we are mostly concerned by the statistical significance of our effects, we have not assessed confidence intervals - they could however have proven informative in Section 5.3. More importantly, we are aware that we reason on the basis of multiple tests due to our setting with multiple units,<sup>52</sup> which might lead to an increase in the number of false rejections. Similarly to Billmeier and Nannicini (2013), we have tried to present our multiple results in an orderly way, and took reasonable precautions in our interpretations.<sup>53</sup> Methodological research on the synthetic control is progressing fast, and theoretical papers offering practicable solutions to this issue (and others) could be adapted to our setting (see notably Abadie and L'hour (2020) and Dube and Zipperer (2015)).

---

<sup>52</sup>When we test multiple variables, a similar issue would appear if we had more emphasized the

<sup>53</sup>For instance, we have not claimed to have identified a result based on 1 significant effect among 9 non-significant effects, as this significant effect could clearly appear at random, even in the absence of a true effect.

## References

- Abadie, A., Diamond, A., & Hainmueller, J. (2010, June). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, *105*(490).
- Abadie, A., Diamond, A., & Hainmueller, J. (2011). Synth: An R package for synthetic control methods in comparative case studies.
- Abadie, A., Diamond, A., & Hainmueller, J. (2014). Synth: Stata module to implement synthetic control methods for comparative case studies.
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495–510.
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *The American Economic Review*, *93*(1), 113–132.
- Abadie, A., & L'hour, J. (2020, october). A penalized synthetic control estimator for disaggregated data. *mimeo*.
- Allen, T., Fichera, E., & Sutton, M. (2016). Can payers use prices to improve quality? Evidence from English hospitals. *Health economics*, *25*(1), 56–70.
- ANAP, & HAS. (2012). *Ensemble pour le developpement de la chirurgie ambulatoire, socle de connaissances*. (Tech. Rep.).
- ANAP, & HAS. (2014). *Day surgery tariffs in France and in other countries : Current situation and future prospects* (Tech. Rep.).
- ATIH. (2014). *Colloque de l'Union des Chirurgiens de France et de la Société française de chirurgie ambulatoire 11 octobre 2014* (Tech. Rep.).
- Bert, T., D'Autume, C., Dupays, S., Hausswalt, P., Perlberg, J., & Weill, M. (2014). *Perspectives du développement de la chirurgie ambulatoire en France* (Tech. Rep.).

- Billmeier, A., & Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3), 983–1001.
- Boadway, R., Marchand, M., & Sato, M. (2004). An optimal contract approach to hospital financing. *Journal of health economics*, 23(1), 85–110.
- Busse, R., Geissler, A., Aaviksoo, A., Cots, F., Häkkinen, U., Kobel, C., . . . others (2013). Diagnosis related groups in europe: moving towards transparency, efficiency, and quality in hospitals? *Bmj*, 346.
- Cash, E., Cash, R., & Dupillet, C. (2011). La réactivité des établissements de santé aux incitations tarifaires. *DREES, Dossier Solidarité & Santé*, 106.
- Cutler, D. M. (1995). The incidence of adverse medical outcomes under prospective payment. *Econometrica*, 63(1), 29–50.
- Dafny, L. S. (2005). How do hospitals respond to price changes? *The American Economic Review*, 95(5), 1525–1547.
- DGOS, & ATIH. (2017). *Le taux de réhospitalisation dans un délai de 1 à 7 jours en MCO* (Tech. Rep.).
- Dube, A., & Zipperer, B. (2015). Pooling multiple case studies using synthetic controls: An application to minimum wage policies.
- Firpo, S., & Possebom, V. (2018). Synthetic control method: Inference, sensitivity analysis and confidence sets. *Journal of Causal Inference*, 6(2).
- Gershlick, B. (2016). Best Practice Tariffs. Country Background Note: United Kingdom (England). *The Health Foundation*.
- Gilman, B. H. (2000). Hospital response to DRG refinements: the impact of multiple reimbursement incentives on inpatient length of stay. *Health economics*, 9(4), 277–294.
- Hafsteinsdottir, E. J. G., & Siciliani, L. (2010). DRG prospective payment systems: refine or not refine? *Health economics*, 19(10), 1226–1239.



- HAS, & ANAP. (2013). *Recommandations organisationnelles de la chirurgie ambulatoire* (Tech. Rep.).
- HAS, & ANAP. (2015). *Outil logiciel permettant l'analyse de l'adéquation des tarifs aux coûts de production de la chirurgie ambulatoire par méthode de microcosting* (Tech. Rep.).
- Heard, K., O'Toole, E., Naimpally, R., & Bressler, L. (2017). Real world challenges to randomization and their solutions. *Boston, MA: Abdul Latif Jameel Poverty Action Lab*.
- Heaton, P., & Helland, E. (2009). Does treatment respond to reimbursement rates? Evidence from trauma care.
- Januleviciute, J., Askildsen, J. E., Kaarboe, O., Siciliani, L., & Sutton, M. (2016). How do hospitals respond to price changes? Evidence from Norway. *Health economics*, 25(5), 620–636.
- Jürges, H., & Köberlein, J. (2015). What explains DRG upcoding in neonatology? the roles of financial incentives and infant health. *Journal of health economics*, 43, 13–26.
- Kaul, A., Klößner, S., Pfeifer, G., & Schieler, M. (2017, July). *Synthetic control methods: Never use all pre-intervention outcomes as economic predictors together with covariates*.
- Kiss, A., Kiss, N., & Váradi, B. (2019). Do Budget Constraints Limit Access to Health Care? Evidence from PCI treatments in Hungary.
- Liang, L.-L. (2015). Do diagnosis-related group-based payments incentivise hospitals to adjust output mix? *Health economics*, 24(4), 454–469.
- Lindrooth, R. C., Bazzoli, G. J., & Clement, J. (2007). The effect of reimbursement on the intensity of hospital services. *Southern Economic Journal*, 575–587.
- Mahieu, A., & Raffy-Pihan, N. (1998). La chirurgie ambulatoire en France bilan et perspectives. *Questions d'économie de la santé*, 2.

- McDonald, R., Zaidi, S., Todd, S., Konteh, F., Hussain, K., Roe, J., ... Sutton, M. (2012). A qualitative and quantitative evaluation of the introduction of best practice tariffs: An evaluation report commissioned by the department of health.
- Milcent, C. (2020). From downcoding to upcoding: Drg based payment in hospitals. *International Journal of Health Economics and Management*, 1–26.
- Newhouse, J. P. (1989). Do unprofitable patients face access problems? *Health Care Financing Review*, 11(2), 33–43.
- OECD. (2017). *Caring for quality in health : Lessons learnt from 15 reviews of health care quality* (Tech. Rep.).
- O’Neill, S., Kreif, N., Grieve, R., Sutton, M., & Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16(1-2), 1–21.
- Papanicolas, I., & McGuire, A. (2015). Do financial incentives trump clinical guidance? hip replacement in England and Scotland. *Journal of Health Economics*, 44, 25–36.
- Sacarny, A. (2018). Adoption and learning across hospitals: The case of a revenue-generating practice. *Journal of health economics*, 60, 142–164.
- Salm, M., & Wuebker, A. (2015). Do hospitals respond to increasing prices by supplying fewer services? *IZA Discussion Paper*, 9229.
- Shigeoka, H., & Fushimi, K. (2014). Supplier-induced demand for newborn treatment: Evidence from japan. *Journal of health economics*, 35, 162–178.
- Shin, E. (2018). Hospital responses to price shocks under the prospective payment system. *Health economics*.
- Siciliani, L. (2009). Paying for performance and motivation crowding out. *Economics letters*, 103(2), 68–71.

- Verzulli, R., Fiorentini, G., Lippi Bruni, M., & Ugolini, C. (2016). Price changes in regulated healthcare markets: Do public hospitals respond and how? *Health economics*.
- Yilmaz, E., & Vuagnat, A. (2015a). Tarification à l'activité et réadmission. *Economie et statistique*, 475(1), 71–87.
- Yilmaz, E., & Vuagnat, A. (2015b). Tarification à l'activité : quel impact sur les réadmissions à l'hôpital ? etude et résultats. *Études et résultats, Drees*(0922).
- Yip, W. C. (1998). Physician response to medicare fee reductions: changes in the volume of coronary artery bypass graft (cabg) surgeries in the medicare and private sectors. *Journal of health economics*, 17(6), 675–699.
- Zhang, H., Wernz, C., & Slonim, A. D. (2016). Aligning incentives in health care: a multiscale decision theory approach. *EURO Journal on Decision Processes*, 4(3-4), 219–244.

## Glossary

- 01C14** Superficial nerve release with the exception of the median carpal tunnel. 25–28, 30, 31, 57, 79, 80, 84, 85
- 01C15** Carpal Tunnel Median Releases. 57, 70, 71
- 02C02** Interventions on the retina. 27, 57, 66, 67, 74, 79, 80
- 02C05** Interventions on the lens with or without vitrectomy. 33–35, 44, 57, 72, 81, 86, 89
- 02C06** Primary interventions on the iris. 27, 30, 57, 66, 79, 80, 84, 85
- 02C08** Other extraocular interventions, age greater than 17 years. 34, 35, 37, 38, 57, 81, 82, 86, 87
- 02C10** Other intraocular procedures for severe conditions. 15, 57
- 02C11** Other intraocular interventions outside severe conditions. 27, 30, 57, 66, 67, 79, 80, 84, 85
- 02C12** Interventions on the lens with trabeculectomy. 15, 57
- 02C13** Interventions on oculomotor muscles, age <18 years. 36–38, 57, 71, 82, 87
- 03C06** Interventions on sinuses and mastoid process, age <18 years. 34, 37, 38, 57, 68, 81, 82, 87
- 03C07** Interventions on sinuses and mastoid process, age greater than 17. 37, 57, 68, 69, 76, 77, 81, 82, 87

- 03C09** Rhinoplasty. 15, 57, 73, 76, 77, 79, 80, 85
- 03C12** Interventions on tonsils and adenoids other than isolated tonsillectomies and / or adenoidectomies, age <18 years. 57, 71
- 03C14** Transtympanic Drains, Age Under 18. 57, 70, 71
- 03C15** Transtympanic drains, age greater than 17 years. 30, 31, 57, 70, 80, 85
- 03C17** Interventions on the mouth. 34–38, 57, 81, 82, 86, 87
- 03C19** Osteotomies of the face. 57, 71, 72, 82
- 03C21** Interventions for protruding ears. 25, 27, 28, 30, 31, 57, 79, 80, 84, 85
- 06C10** Restorative procedures for hernias and eventrations, age under 18. 33–38, 57, 81, 82, 86, 87
- 06C12** Restorative procedures for inguinal and crural hernias, age greater than 17. 25–28, 30, 31, 57, 79, 80, 84, 85
- 06C14** Interventions on the rectum and anus other than rectal resections. 33–35, 37, 38, 44, 57, 81, 82, 86–88
- 06C15** Other interventions on the gastrointestinal tract outside laparotomies. 30, 57, 67, 80, 85
- 06C19** Hemorrhoidectomies. 37, 38, 57, 76, 77, 81, 82, 86, 87
- 06C25** Restorative procedures for hernias except for inguinal hernias, crural, age greater than 17. 15, 57, 74, 75, 81, 82
- 07C11** Biliary Derivations. 57, 74

- 07C14** Cholecystectomies without exploration of the main bile duct except acute diseases. 26, 27, 57, 66, 74, 79, 80, 84
- 08C04** Interventions on the hip and femur, age <18 years. 15, 30, 57, 67, 80, 85
- 08C12** Osteoarticular Biopsies. 27, 28, 57, 79, 84
- 08C13** Localized bone resections and / or removal of internal fixation material at the hip and femur. 15, 57
- 08C14** Localized bone resections and / or removal of internal fixation material at a location other than the hip and femur. 27, 30, 57, 66, 67, 79, 80, 84, 85
- 08C20** Skin grafts for musculoskeletal or connective tissue disease. 27, 30, 57, 66, 67, 79, 80, 85
- 08C28** Maxillofacial Interventions. 30, 31, 57, 76, 77, 79, 80, 84, 85
- 08C31** Interventions on the leg, age under 18. 27, 30, 57, 66, 67, 79, 80, 84, 85
- 08C33** Interventions on ankle and hindfoot with the exception of fractures. 27, 30, 57, 66, 67, 79, 80, 84, 85
- 08C34** Interventions on Arthroscopic Cross Ligaments. 37, 38, 57, 71, 81, 82, 87
- 08C35** Interventions on the arm, elbow and shoulder. 37, 57, 69, 82, 87
- 08C38** Other knee arthroscopy. 33–35, 57, 72, 81, 86
- 08C40** Arthroscopies from other locations. 27, 57, 66, 76, 77, 79, 82, 84, 87
- 08C43** Non-minor interventions on the hand. 37, 57, 69, 82, 87
- 08C44** Other interventions on the hand. 26–28, 30, 31, 57, 79, 80, 84, 85
- 08C45** Arthroscopic meniscectomy. 33–35, 57, 72, 81, 86

- 08C46** Other interventions on soft tissues. 34, 35, 37, 38, 57, 81, 82, 86, 87
- 08C54** Interventions on the knee for non-traumatic conditions. 34, 35, 37, 38, 57, 81, 82, 86, 87
- 08C56** Interventions for osteoarticular infections. 15, 57
- 08C58** Arthroscopies of the shoulder. 26, 27, 30, 57, 66, 67, 79, 80, 84, 85
- 08C59** wrist tenosynovectomies. 37, 38, 57, 82, 87
- 08C60** Interventions on the wrist other than tenosynovectomies. 33–35, 37, 38, 57, 81, 82, 86, 87
- 09C07** Local biopsies and excisions for nonmalignant breast conditions. 25–28, 30, 31, 57, 79, 80, 84, 85
- 09C08** Interventions on the anal and perianal region. 25, 27, 28, 30, 31, 57, 79, 80, 84, 85
- 09C13** Interventions for anogenital condyloma. 25, 27, 28, 30, 31, 57, 79, 80, 84, 85
- 09C14** Lymphonodal dissection for skin, subcutaneous tissue or breast disease. 37, 57, 69, 82, 87
- 09C15** Interventions on skin, subcutaneous tissue or breasts for traumatic injuries. 57
- 10C05** Interventions on parathyroid glands. 27, 57, 66, 74, 79, 80, 84
- 10C07** Interventions on the thyroglossal tract. 30, 57, 67, 73, 79, 80, 85
- 10C10** Other interventions for obesity. 30, 57, 67, 80, 85

- 10C11** Thyroid Interventions for Malignant Tumors. 57, 73, 74, 79
- 10C12** Thyroid Interventions for Non-malignant Conditions. 26, 30, 57, 67, 73, 79, 80, 85
- 11C07** Interventions on the urethra, age greater than 17. 34, 37, 57, 69, 81, 82, 86, 87
- 12C04** Transurethral Prostatectomies. 15, 57, 74, 76, 77, 79, 84
- 12C05** Interventions on testes for malignant tumors. 26, 27, 30, 31, 57, 66, 79, 80, 84, 85
- 12C07** Interventions on testes for non-malignant conditions, age greater than 17. 57, 76, 77, 81, 82, 86, 87
- 12C08** Circumcision. 30, 31, 57, 70, 80, 85
- 12C13** Sterilization and vasoplasty. 30, 31, 57, 70, 80, 85
- 13C04** Restorative Interventions on the Female Genitalia. 27, 30, 57, 66, 67, 79, 80, 84, 85
- 13C05** Interventions on the uteroannexial system for malignant tumors. 34, 37, 57, 68, 69, 81, 82, 86, 87
- 13C06** Tubal interruptions. 34, 57, 68, 75, 81, 82
- 13C07** Interventions on the uteroannexial system for nonmalignant conditions, other than tubal interruptions. 57, 75, 82
- 13C08** Interventions on the vulva, vagina or cervix. 25–28, 30, 31, 57, 79, 80, 84, 85
- 13C09** Laparoscopies or diagnostic laparoscopies. 57, 76, 77, 81, 86



- 13C10** Tubal ligation by laparoscopy or laparoscopy. 33–35, 57, 76, 77, 81, 82, 86, 87
- 13C11** Dilatations and curettage, conizations for malignant tumors. 34, 35, 57, 72, 81, 86
- 13C12** Dilation and curettage, conations for non-malignant conditions. 25–28, 30, 31, 57, 79, 80, 84, 85
- 13C16** Oocyte sampling, outpatient. 57, 71, 72
- 13C17** Cervicocystopexy. 30, 31, 57, 76, 77, 79, 80, 84, 85
- 13C19** Interventions for infertility or reproductive care reasons. 34–38, 57, 81, 82, 86, 87
- 13C20** Exeresis or destruction of cervical lesions except conizations. 30, 31, 57, 70, 80, 85
- 21C04** Interventions on the hand or wrist due to injury. 37, 38, 57, 74, 81, 82, 87
- 21C06** Skin grafts or wound sites for lesions other than burns. 15, 27, 30, 57, 66, 67, 79, 80, 85

# Appendices

## A Characteristics of the data base

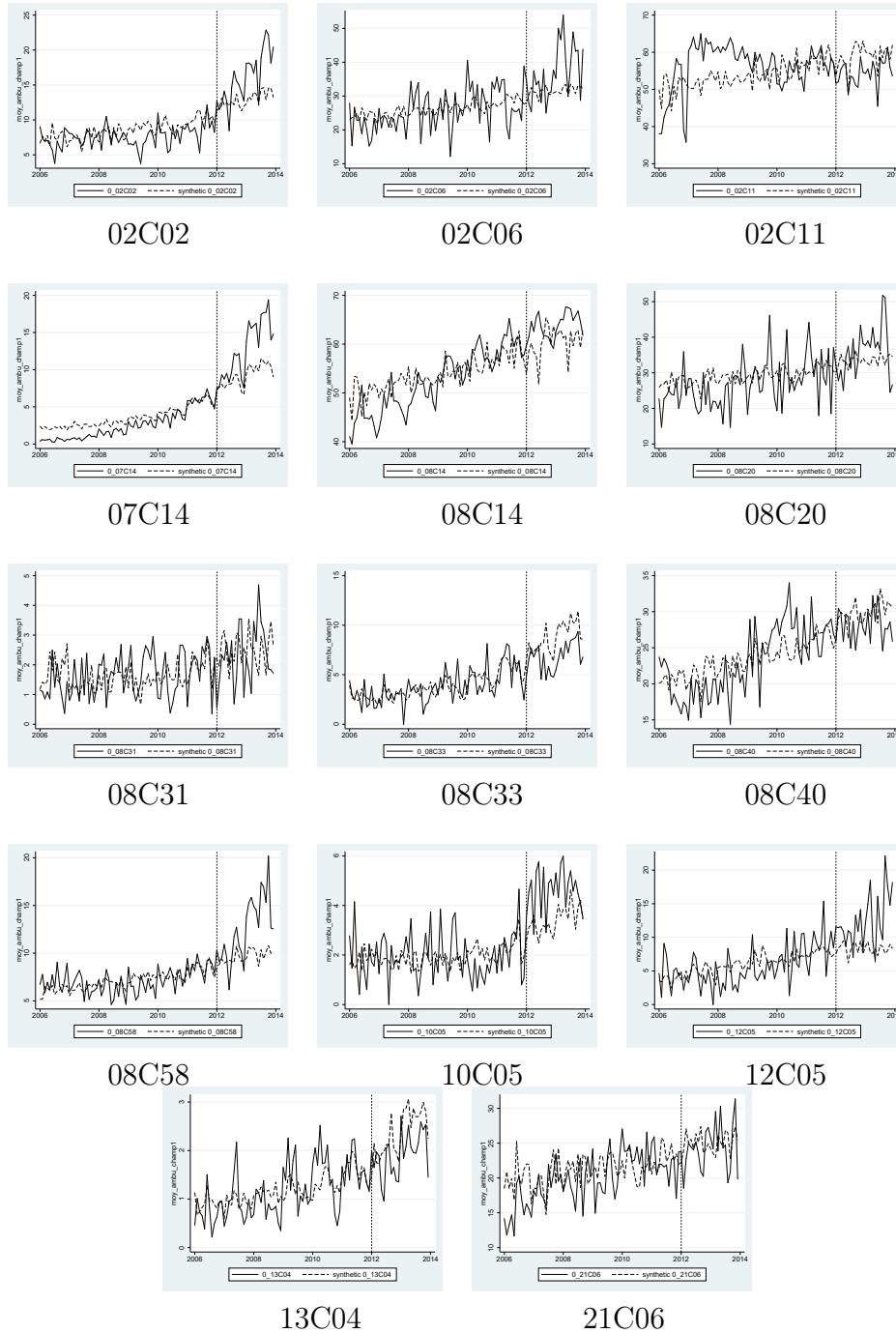
Table 10: The characteristics of the data base

	2006	2007	2008	2009	2010	2011	2012	2013
<b>Severity of the stay (in %)</b>								
1	56	56	54	50	48	46	44	42
2	7	8	8	9	9	9	9	9
3	2	3	3	3	4	4	4	4
4	1	1	1	1	1	2	2	2
practice of same-day surgery	33	32	33	36	38	40	41	43
<b>Financial weight of the hospital (en %)</b>								
With the tariffs of the private-for-profit sector	61	60	59	59	58	58	58	58
With the tariffs of the public sector	39	40	41	41	42	42	42	42
<b>Patients characteristics</b>								
Men (in %)	47	48	48	48	48	48	48	48
Mean age (in years)	51.5	51.8	52.2	52.6	52.9	52.9	53.2	53.6
Mean length of stay (in days)	4.0	4.0	3.9	3.7	3.6	3.5	3.4	3.3
Mean associated diagnoses	1.4	1.4	1.4	1.5	1.7	1.8	1.9	1.9
Mean acts	4.3	4.6	5.0	5.2	5.4	5.3	5.4	5.5
Total number of stays	5,035,190	4,951,199	4,978,072	5,078,551	5,136,940	5,288,089	5,372,189	5,406,323

**Field:** Surgical stays (save surgical operations on pathological pregnancies and births) in France from 2006 to 2014.  
**Source:** ATIH, PMSI-MCO (version v11e of the French DRG classification), treatment by the authors.

## **B Graphs for the DRGs incentivized starting from 2012**

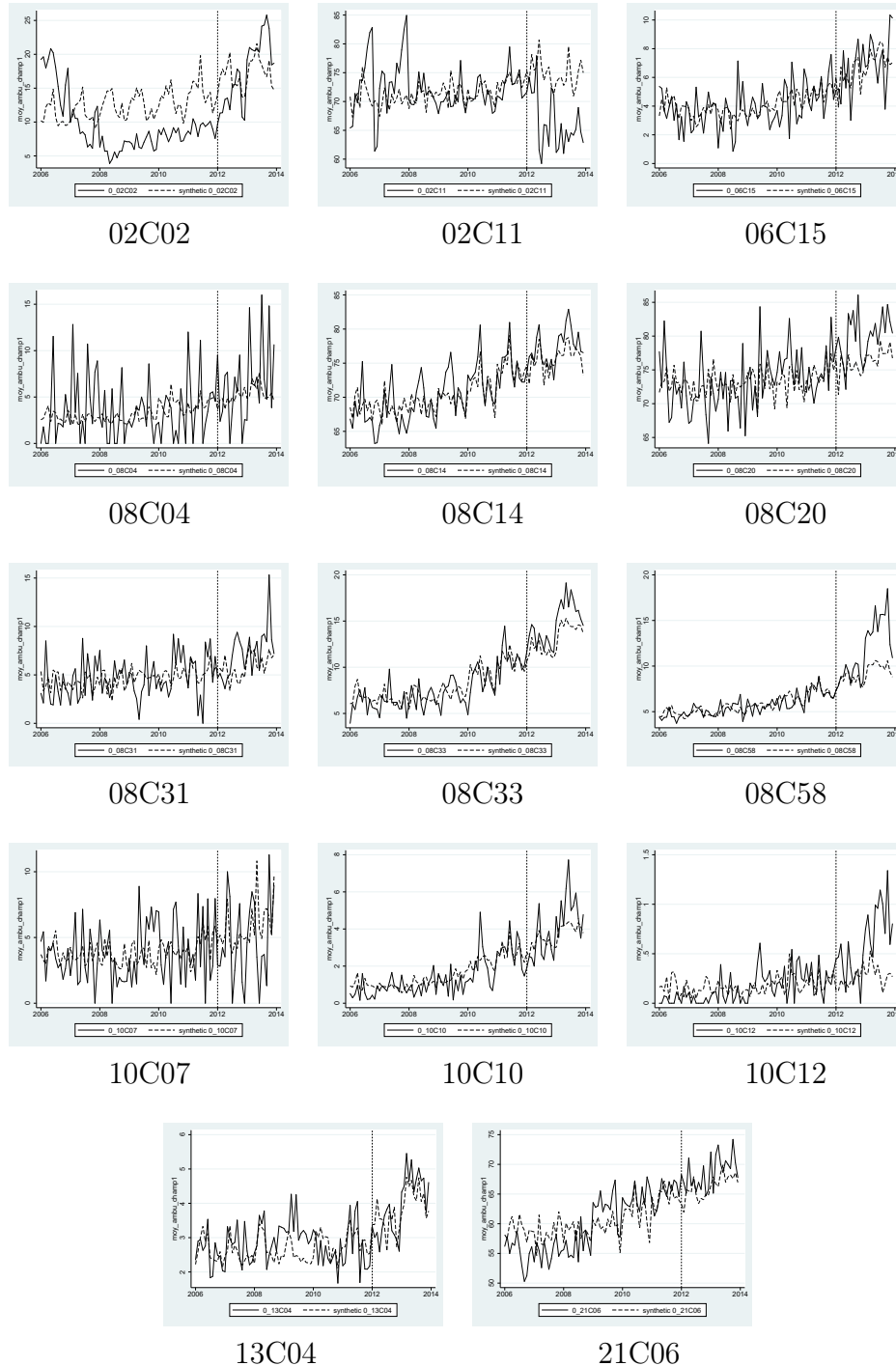
Figure 6: Rates of same-day surgery for DRGs with the single tariff policy in 2012 (public sector)



**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.

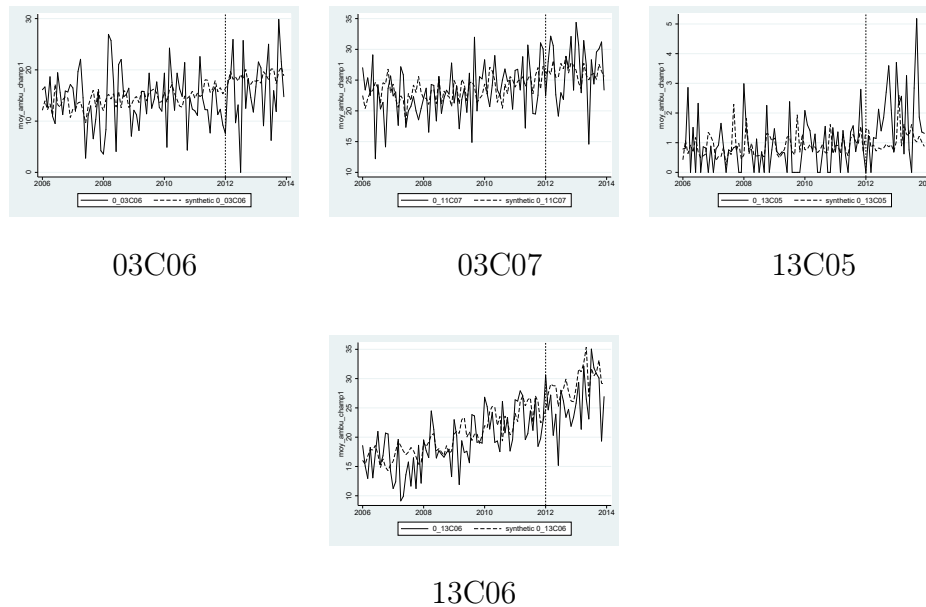
Figure 7: Rates of same-day surgery for DRGs with the single tariff policy in 2012 (private sector)



**Reading :** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.

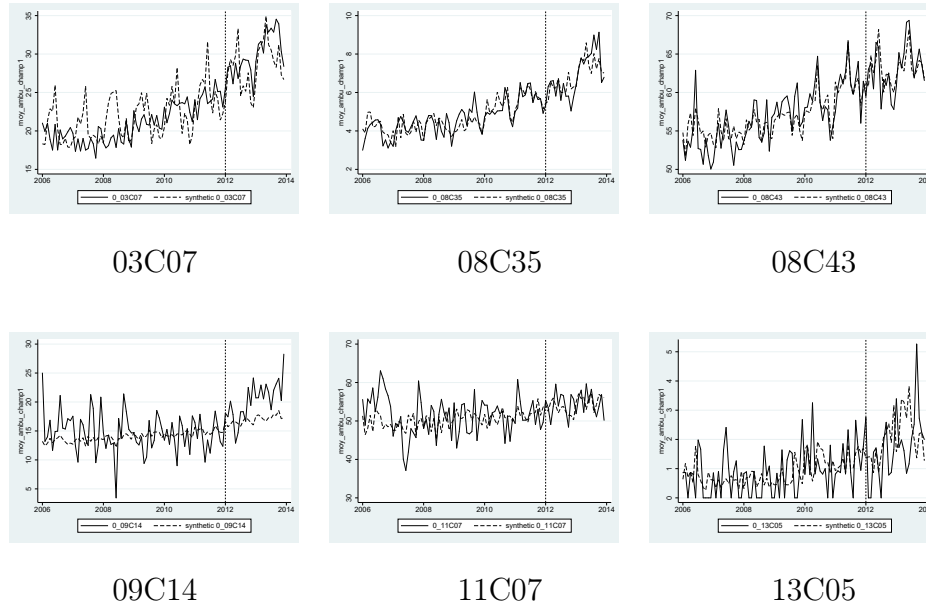
Figure 8: Rates of same-day surgery for DRGs with only the prior approval policy in 2012 (public sector)



**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.

Figure 9: Rates of same-day surgery for DRGs with only the prior approval policy in 2012 (private sector)



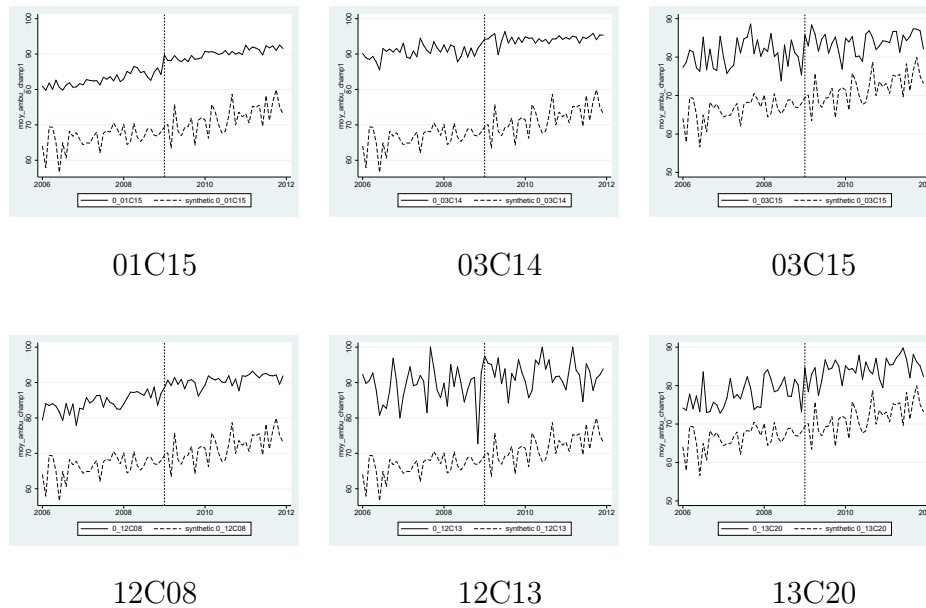
**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.



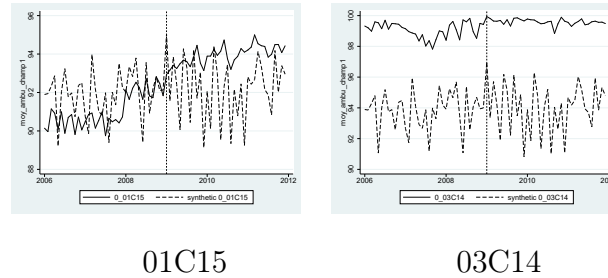
## C DRGs for which no reasonable synthetic control was found.

Figure 10: Rates of same-day surgery for DRGs with the single tariff policy in 2009 (public sector)



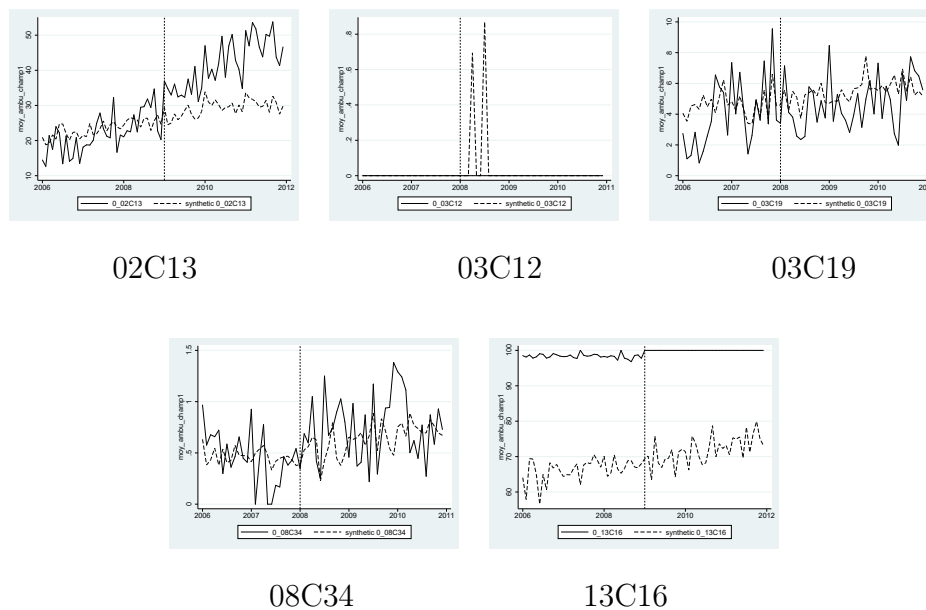
**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied. **Source :** ATIH, PMSI-MCO, treatments by the authors.

Figure 11: Rates of same-day surgery for DRGs with the single tariff policy in 2009 (private sector)



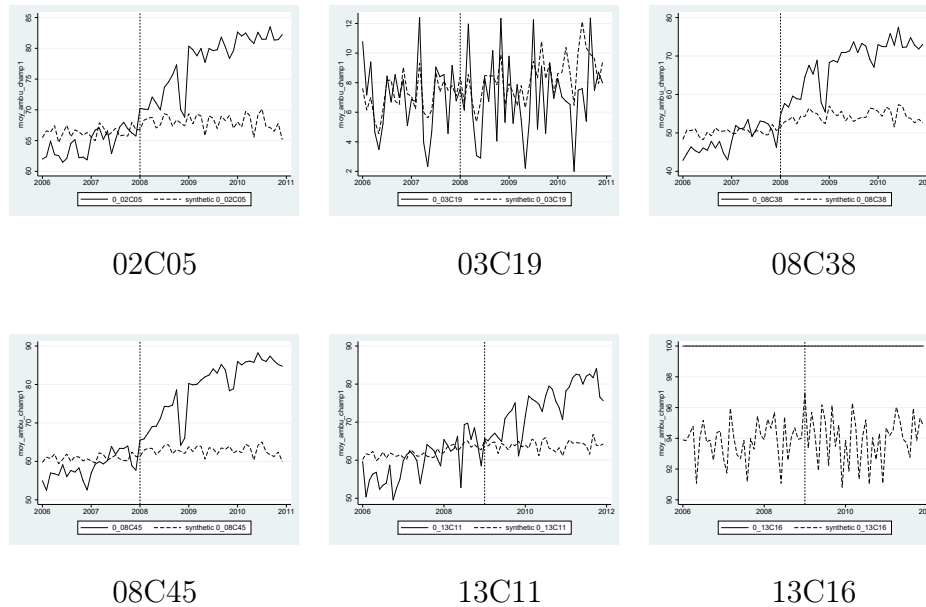
**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied. **Source :** ATIH, PMSI-MCO, treatments by the authors.

Figure 12: Rates of same-day surgery for DRGs with only the prior approval policy in 2008 or in 2009 (public sector)



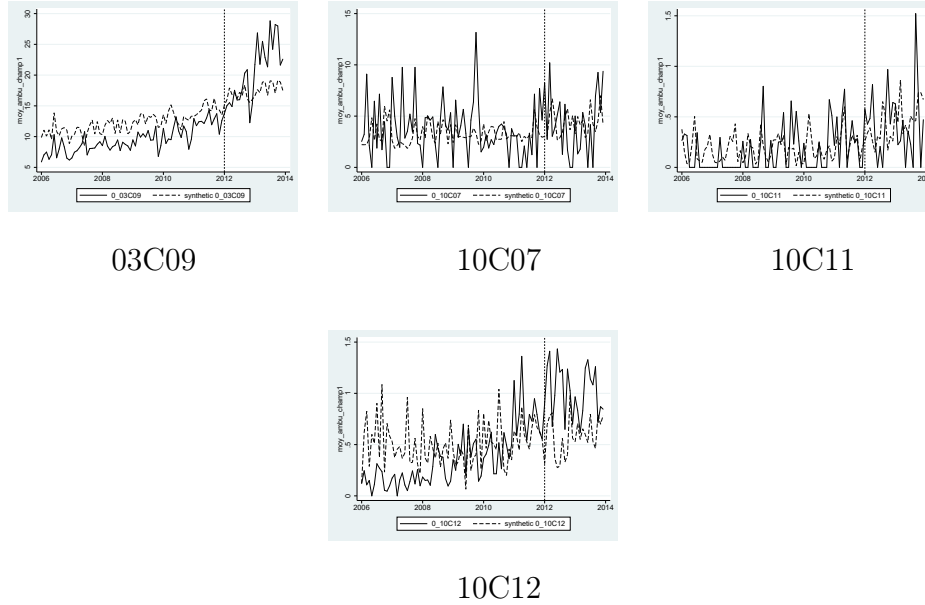
**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied. **Source :** ATIH, PMSI-MCO, treatments by the authors.

Figure 13: Rates of same-day surgery for DRGs with only the prior approval policy in 2008 or in 2009 (private sector)



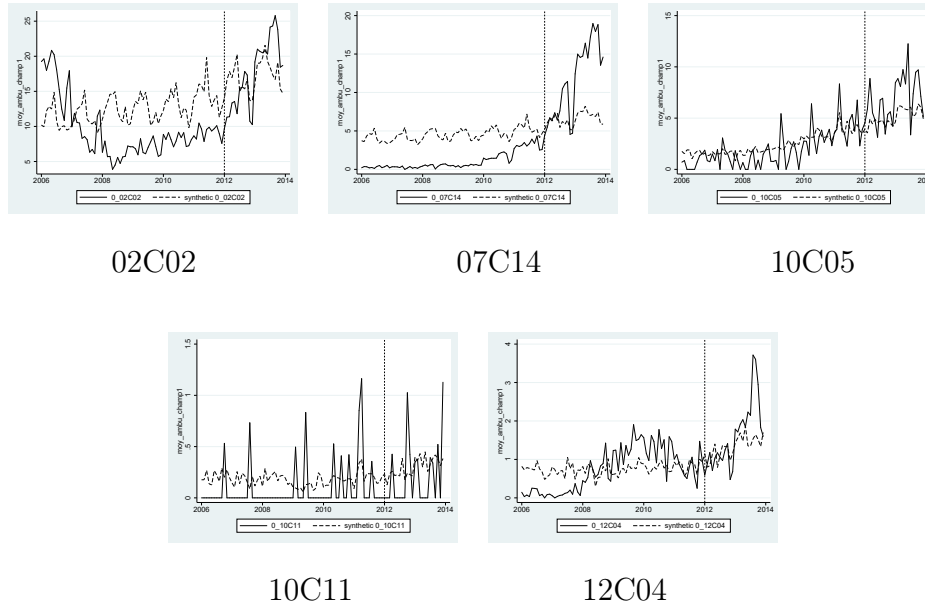
**Reading :** for every studied DRG in 2009, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied. **Source :** ATIH, PMSI-MCO, treatments by the authors.

Figure 14: Rates of same-day surgery for DRGs with the single tariff policy in 2012 (public sector)



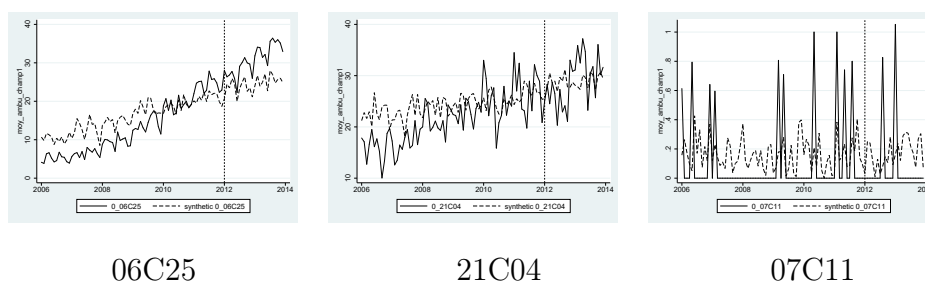
**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied. **Source:** ATIH, PMSI-MCO, treatments by the authors.

Figure 15: Rates of same-day surgery for DRGs with the single tariff policy in 2012 (private sector)



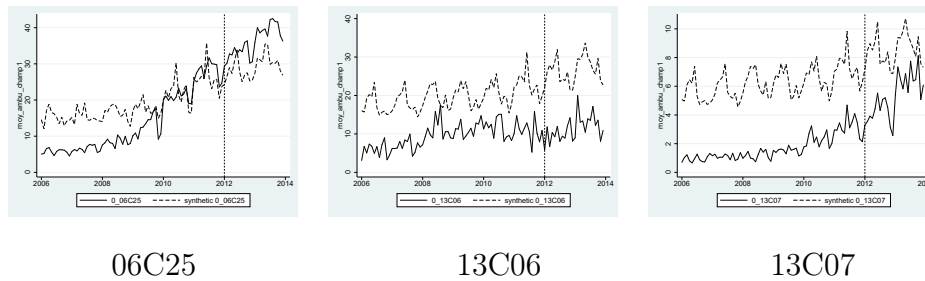
**Reading :** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied. **Source:** ATIH, PMSI-MCO, treatments by the authors.

Figure 16: Rates of same-day surgery for DRGs with only the prior approval policy in 2012 (public sector)



**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied. **Source:** ATIH, PMSI-MCO, treatments by the authors.

Figure 17: Rates of same-day surgery for DRGs with only the prior approval policy in 2012 (private sector)



**Reading:** for every studied DRG in 2012, its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied. **Source:** ATIH, PMSI-MCO, treatments by the authors.

Table 11: Estimates with a placebo effect in  $T_{d-1}$  that is statistically significant

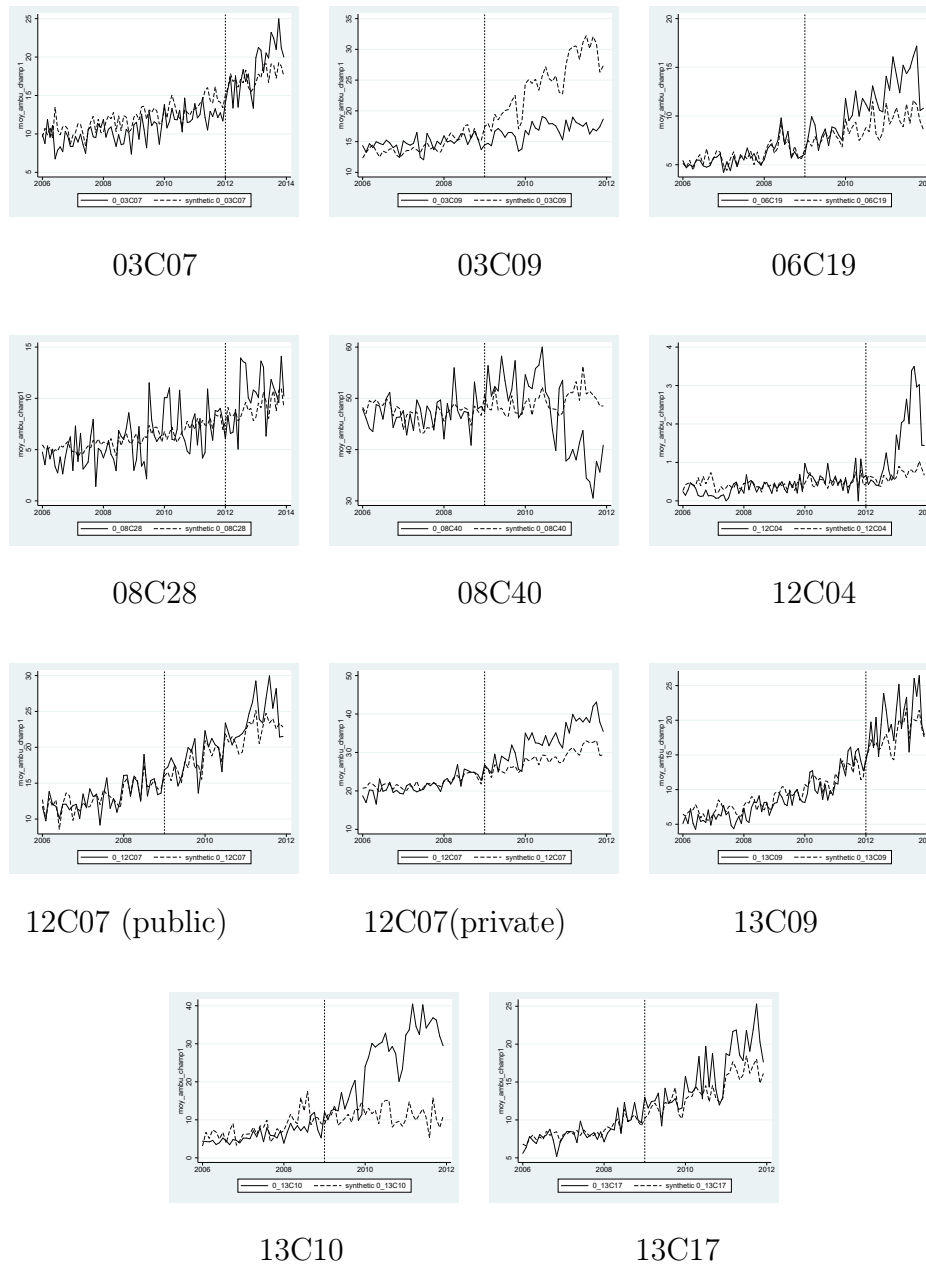
DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Sector	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
13C17	2009	yes	no	public	2.03**	0.67	1.82**	3.73**	0	.2	1*
12C04	2012	yes	no	public	0.81**	0.1	1.52**	-	0	0	-0.24**
08C28	2012	yes	weak	public	1.12	0.98	1.51	-	2.7	-1.7	-2.24*
03C09	2009	yes	no	private	-7.73**	-3.74**	-7.74**	-12.7**	1.6	-.3	-2.66**
06C19	2009	no	weak	public	2.29**	0.49	2.48**	4.15**	.2	-.1	0.72*
12C07	2009	no	weak	public	1.18	0.59*	0.99	2.33*	.4	-.5	1.06*
13C09	2012	no	weak	public	2.21	2.35**	2.3	-	-.9	1.4	1.97**
03C07	2012	no	weak	public	1.07	-0.85	2.77	-	9.2	7.5	-3.35*
08C40	2009	no	weak	private	-2.19*	3.68	3.06	-12.74*	.4	0	0.44**
12C07	2009	no	weak	private	4.84*	1.97	4.95*	7.87*	.2	-.3	2.27**
13C10	2009	no	sizeable	private	14.09**	2.21	15.88**	25.16**	7	0	-3.12*

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column signals when the policy starts. "no" (resp. "weak" / "sizeable") loose prior approval policy means that less than 10% of the stays of the DRG (resp. between 10% and 70% / more than 70%) that benefited from the loose prior approval policy. DRGs are sorted in increasing order of  $T_d$  and of this latter variable.

Figure 18: Rates of same-day surgery for DRGs with a placebo effect that is statistically significant



**Reading :** for every studied DRG in 2012 (for instance), its rate of same-day surgery is in straight line and its synthetic control is in dashed line. The vertical line in March 2012 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.



## D Results tables when the data-driven matrix $V$ is used

These data show the results when the data-driven matrix  $V$  is used in the synthetic control procedure; it assigns almost a zero weighting in the optimisation procedure to the low- and high- bound same-day prospects defined by the learned medical societies.

As an *ex-post* consequence, the synthetic controls often have often much lower same-day prospects than the incentivized DRGs, which overestimates the impacts of the policies. In particular, this would cause several significant effects to appear, notably in the private sector.

Table 12: Estimates in the public sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo effect in $T_d - 1$
08C12	2009	yes	no	18.2	1.72	0.72	2.81	1.49	-6.7	16.9	3.62
03C21	2009	yes	no	38.2	3.83	0.8	5.1	5.81	-9.7	-5.6	-2.34
13C17	2009	yes	no	12.3	2.06**	0.73	1.63*	3.85**	4.5	10.2	1.04*
09C08	2009	yes	no	15.9	5.28***	2.67***	5.95***	7.91***	4.5	1.7	-0.1
13C08	2009	yes	no	40.8	3.82	2.89**	4.23*	5.2	.2	4.2	2.15
09C13	2009	yes	no	49.3	18.88**	7.62	23.72**	27.1*	13.7	12.6	2.54
08C44	2009	yes	weak	51.8	2	1.96	2.19	2.26	8.9	8.4	1.72
01C14	2009	yes	strong	55.5	4.99	3.57	5.34	6.9	5.4	4.7	3.97
13C12	2009	yes	strong	66.7	9.61**	6.47**	10.85**	12.87**	9.1	10.6	3.33
06C12	2009	yes	strong	18.2	8.48***	3.44***	9.92***	13.07***	21.5	21.3	2.81**
09C07	2009	yes	strong	33.7	7.98**	2.46	7.88**	14.37**	17.2	31.2	1.59
02C06	2012	yes	no	32.6	8.3	2.3	14.85	-	23.6	25.3	0.33
10C12	2012	yes	no	1	0.21	0.34*	0.08	-	-1.5	-2.3	0.36**
10C07	2012	yes	no	3.9	-0.71	-0.62	-1.23	-	-2	-2.8	-1.32
10C05	2012	yes	no	4.4	1.09	1.36	0.97	-	2.6	1.7	0.27
02C11	2012	yes	no	53.9	-7.42	-7.6	-8.15	-	-10.3	-11.6	-2.45
10C11	2012	yes	no	.5	0.05	0.14	0	-	0	.3	0.15
12C04	2012	yes	no	.6	0.73**	0.03	1.43**	-	.4	0	-0.25**
08C20	2012	yes	no	33.5	4.19	1.38	7.49	-	26.6	24.2	3.35
08C31	2012	yes	no	2	-0.14	-0.44	-0.04	-	.3	5.8	0.07
08C33	2012	yes	no	6	-1.23	-1.08	-1.48	-	2.1	.9	0.46
02C02	2012	yes	no	13.5	5.33**	3.27*	7.7**	-	11.9	25.4	0.29
13C04	2012	yes	no	1.7	-0.26	-0.22	-0.32	-	1.7	2.4	-0.1
08C14	2012	yes	no	62.2	4.29	4.14	4.22	-	-1.3	-7.2	3.82
21C06	2012	yes	no	23.5	0.71	-0.34	2.08	-	33.1	29.5	-1.57
12C05	2012	yes	no	10.8	1.38	0.59	2.11	-	-4.4	-6.3	1.83
08C28	2012	yes	weak	9.4	0.91	0.87	1.17	-	9.4	3.8	-0.4
08C40	2012	yes	weak	27.8	-1.7	-0.88	-2.67	-	8.7	5.5	-1.78
03C09	2012	yes	weak	16.2	4.34**	1.34	7.47**	-	20.7	20.4	0.05
08C58	2012	yes	weak	10	2.8*	0.58	4.92*	-	-6	-9	-0.01
07C14	2012	yes	strong	9.5	5.1**	3.01**	7.36**	-	4.5	6.7	1.69**

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column  $T_d$  signals when the policy starts. We highlight in bold differences in potential rates higher than the highest estimated effect (and of same sign).

Table 13: Estimates in the private sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
12C08	2009	yes	no	96.2	2.62	1.95	3.81	3.03	6.5	4.5	1.22
03C15	2009	yes	no	94.8	0.77	0.4	1.69	0.69	4.5	3.5	0.71
03C21	2009	yes	no	40.2	3.94	2.78	3.68	5.35	20.7	23.1	3.48
03C09	2009	yes	no	15.4	-4.29**	-2.34*	-4.07**	-7.02**	10.4	8.9	-0.09
09C08	2009	yes	no	22.9	-0.72	0.75*	-0.49	-2.33	.5	-2.8	0.58
13C17	2009	yes	no	16.8	3.08*	1.75	2.84	4.84*	1.5	-2.8	1.4
08C28	2009	yes	no	37	-1.42	-1.54	-1.21	-2.68	-8.9	-16.4	-0.75
09C13	2009	yes	no	79.9	-0.4	-0.23	-1.25	0.1	3.7	5.1	4.39
13C08	2009	yes	no	59.7	4.67	2.35	6.08	6.07	1.3	5.4	1.91
12C05	2009	yes	weak	5.8	3.49*	0.59	2.84	7.13*	6.3	7.3	0.35
08C44	2009	yes	weak	83.4	1.39	0.54	1.91	2.04	0	1.7	1.22
13C20	2009	yes	weak	92	3.19	1.7	4.64	4.41	9.2	12.7	1.9
12C13	2009	yes	strong	97.4	3.63	3.35	4.8	3.8	6.9	4.9	-0.13
13C12	2009	yes	strong	77.9	7.89**	4.22*	9.23**	11.12*	13.5	14.8	2.63*
01C14	2009	yes	strong	68.9	6.98*	4.96**	7.86**	9.1*	5.4	-2.2	3.02
06C12	2009	yes	strong	15.9	14.73***	4.66***	17.63***	23.46***	26	26	1.74**
09C07	2009	yes	strong	46.3	10.77**	5.31*	11.56**	16.89*	13.5	20.5	4.32*
10C12	2012	yes	no	.3	0.35*	0.13	0.56*	-	0	-.2	0
10C05	2012	yes	no	5.5	1.24	0.62	1.79	-	.2	-2.6	0.89
10C10	2012	yes	no	2.8	0.64	0.03	1.23	-	1.4	.7	0.77
02C06	2012	yes	no	76	2.64	1.09	4.08	-	-12.9	-12.4	2.49
10C07	2012	yes	no	3.8	-0.46	-0.45	-0.87	-	2.1	3.3	-0.42
02C02	2012	yes	no	13.2	6.01	2.59	9.46	-	16.8	32	-0.86
08C33	2012	yes	no	12.7	2.58	1.76	3.41	-	4.2	1.4	1.54
02C11	2012	yes	no	69.9	-9.43	-7.85	-11.06	-	1.5	-1.9	-0.89
08C20	2012	yes	no	79	3.9	4.46	3.4	-	-.4	.1	-3.16
08C04	2012	yes	no	5.3	0.23	-0.87	0.9	-	-3.5	-1.4	1.29
13C04	2012	yes	no	3.2	0.27	-0.14	0.66	-	.6	.7	-0.23
08C31	2012	yes	no	6	0.94	0.85	0.92	-	.4	6.4	-0.63
21C06	2012	yes	weak	66.9	2.12	1.13	3.28	-	5.4	.5	-0.36
06C15	2012	yes	weak	6.2	0.11	0.16	0.19	-	.5	.1	1.09
08C14	2012	yes	weak	75.6	1.6	0.95	2.09	-	-.3	-4.2	-0.91
08C58	2012	yes	weak	8.8	2.91**	0.73	5.13**	-	2.2	3.2	0.94
07C14	2012	yes	strong	7.2	7.74**	4.08**	11.54**	-	9.8	15.2	1.5**

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column  $T_d$  signals when the policy starts. We highlight in bold differences in potential rates higher than the highest estimated effect (and of same sign).

Table 14: Estimates in the public sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo effect in $T_d - 1$
08C54	2008	no	weak	9.4	3.19	1.92	3.71	4.06	9.2	8.5	-1.86
03C17	2008	no	weak	42.7	2.7	-0.16	3.47	4.94	18	17.2	-2.69
08C34	2008	no	weak	.7	0.24	0.24	0.26	0.33	4.3	13	-0.2*
08C38	2008	no	strong	44.7	13.83*	7.76*	15.45*	20.13*	26.3	29.3	1.04
02C05	2008	no	strong	65.3	16.87*	9.15*	20.46**	23.2*	19.6	20.6	5.31
08C45	2008	no	strong	60.1	16.98**	7.94**	19.77**	25.22**	25.5	31.3	0.03
06C19	2009	no	weak	8	2.48**	0.56	2.69**	4.45**	2.1	-5	0.78*
06C10	2009	no	weak	52.2	6.98	5.71*	9.17*	7.49	1.4	4.9	1.66
08C60	2009	no	weak	29	4.63*	2.88	5.07*	7.13	12.6	9.5	0.66
02C08	2009	no	weak	61.7	2.73	3.16	3.21	2.64	4.4	4.6	0.66
08C46	2009	no	weak	54.1	2.07	1.9	2.57	2.35	3	3.7	1.98
06C14	2009	no	weak	12.7	1.77**	0.56**	1.95**	3.03**	2.2	-2.9	0.41
13C19	2009	no	weak	11.7	2.21	0.13	1.5	5.2	5.8	24.2	1.09
12C07	2009	no	weak	17.5	0.83	0.29	0.62	1.9	-1.7	-4.2	2.13
13C11	2009	no	strong	64.8	8.76*	4.82	8.96*	13.31*	9.3	11.5	2.91
13C10	2009	no	strong	25	11.94**	4.86*	12.76**	19.43**	18	43.4	2.1
21C04	2012	no	weak	26	1.64	-0.38	4.05	-	25	21.8	1.96
13C05	2012	no	weak	1.5	0.69	0.51	0.85	-	.6	2.6	0.21
13C09	2012	no	weak	18.4	2.03	2.12*	2.16	-	-.9	8.1	2.72**
13C06	2012	no	weak	24.4	-1.35	-1.77	-1.02	-	15.3	47.6	-0.18
11C07	2012	no	weak	25.8	0.41	-0.93	1.38	-	8.9	7.1	-0.31
03C06	2012	no	weak	15.2	-2.65	-2.89	-2.59	-	13.3	22.1	-3.2
03C07	2012	no	weak	15.8	4.32**	1.97*	6.69**	-	15.7	26.2	-1.24
06C25	2012	no	strong	28.1	10.62	7.84	13.94	-	6	7.3	7.29*

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column  $T_d$  signals when the policy starts. We highlight in bold differences in potential rates higher than the highest estimated effect (and of same sign).

Table 15: Estimates in the private sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
08C34	2008	no	weak	.3	-0.39**	-0.18**	-0.44**	-0.58**	3.8	12.7	0.01
03C06	2008	no	weak	38.4	1.37	2.76	0.5	1.19	5.5	11.5	5.18
08C54	2008	no	weak	16.1	4.63*	1.86	4.39	8.26*	12.8	12.5	3.2
03C17	2008	no	weak	70.9	7.01**	3.08**	8.34**	10.31*	17.7	17.1	-1.12
03C19	2008	no	weak	6.5	-0.34	-0.56	-0.5	-0.05	8.3	8.1	-0.68
08C40	2009	no	weak	51.3	-1.28	4.07	3.89	-11.22*	4.4	5	2.12
06C19	2009	no	weak	9.7	0.34	-0.01	0.54	0.6	1.3	-3.3	-1.15
06C14	2009	no	weak	23.3	-2.18**	-0.38	-2.33**	-3.91*	.7	-3.9	0.69*
08C59	2009	no	weak	84.4	0.5	0.04	1.38	0.79	6.1	5.1	2.02
21C04	2009	no	weak	72	2.52	1.49	3.97	2.82	-1	.7	-0.57
02C08	2009	no	weak	77.3	-0.18	-0.48	-0.28	0.49	9	9.1	0.77
08C46	2009	no	weak	76	-0.08	-0.81	0.4	0.48	.4	3.9	1.49
06C10	2009	no	weak	55.4	9.37*	1.86	11.75*	15.11*	6.5	2	5.22*
02C13	2009	no	weak	36.5	10.18	4.84	9.25	17.6	20.8	4.4	5.66*
13C19	2009	no	weak	7.1	3.09*	0.38	3.25*	5.95*	9.3	26.8	2.09*
08C60	2009	no	weak	53.5	0.91	0.35	1.61	1.11	4.5	1.5	1.27
12C07	2009	no	weak	27.4	3.36	1.67	3.5	5.17	-1.6	-4.6	2.91**
13C10	2009	no	strong	13.4	10.31**	3.26**	12.47**	16.13**	22.7	40.2	1.76**
13C07	2012	no	weak	4	1.18*	0.53*	1.76*	-	3.3	28.4	0.55*
13C06	2012	no	weak	9.7	-1.78	-2.59	-1.19	-	31.3	57.3	-2.08
13C05	2012	no	weak	1.3	-0.05	-0.31	0.13	-	.4	1.7	-0.05
09C14	2012	no	weak	18.6	2.93	1.61	5	-	5	5.8	-1.69
08C35	2012	no	weak	5.9	-0.17	-0.33	-0.07	-	-4	1.4	0.31
08C43	2012	no	weak	61.2	-0.56	-1.72	0.42	-	1.5	-1	0.36
03C07	2012	no	weak	27.7	2.43	1.73	3.03	-	12.8	23.3	-0.48
11C07	2012	no	weak	54.6	0.89	1.89	0.29	-	-3.1	-5.7	-1.77
06C25	2012	no	weak	32.6	8.83	7.25	10.85	-	14.1	17	6.12

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column  $T_d$  signals when the policy starts. We highlight in bold differences in potential rates higher than the highest estimated effect (and of same sign).

**E Result tables when a weight of 50% is attributed  
to the two control variables**

Table 16: Estimates in the public sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo effect in $T_d - 1$
08C12	2009	yes	no	18.2	3.26	1.83	4.06	4.02	-6	.7	2.57
03C21	2009	yes	no	38.2	5.78	1.65	7.72*	8.35	-6	.3	-1.66
13C17	2009	yes	no	12.3	2.06**	0.71	1.85**	3.76**	0	.1	1.02*
09C08	2009	yes	no	15.9	3.79**	2.22**	3.79**	5.83**	.5	-.4	-0.42
13C08	2009	yes	no	40.8	3.53**	2.57**	3.77**	5.09*	-.2	.4	1.35
09C13	2009	yes	no	49.3	17.06**	6.25	21.76**	24.58**	5.2	4.8	-0.24
08C44	2009	yes	weak	51.8	0.61	1.17	0.63	0.26	1.2	1.1	0.15
01C14	2009	yes	sizeable	55.5	4.23	3	4.28	6.08*	.6	.1	3.26
13C12	2009	yes	sizeable	66.7	6.68**	4.76**	7.14**	9.03*	2.1	1.6	1.87
06C12	2009	yes	sizeable	18.2	11.32*	3	12.81*	19.35**	8.9	5	-3.29
09C07	2009	yes	sizeable	33.7	7.58	2.9	6.9	13.75*	8.4	8.1	-0.95
02C06	2012	yes	no	32.6	6.29	1.88	11.2	-	5	2.7	-1.66
10C05	2012	yes	no	4.4	1.3	1.43*	1.32	-	.4	-.2	0.29
02C11	2012	yes	no	53.9	-2.78	-3.06	-3.06	-	-5.1	-4.8	2.11
12C04	2012	yes	no	.6	0.8**	0.09	1.52**	-	0	0	-0.23**
08C31	2012	yes	no	2	-0.19	-0.4	-0.12	-	0	0	0.15
08C33	2012	yes	no	6	-2.25	-1.94	-2.72	-	.7	-.5	-0.53
13C04	2012	yes	no	1.7	-0.49	-0.37	-0.66	-	.1	-.1	-0.25
08C14	2012	yes	no	62.2	3.42	2.81	3.65	-	.1	-.4	2.41
12C05	2012	yes	no	10.8	4.24	2.54	6	-	-.7	.2	1.99
08C28	2012	yes	weak	9.4	0.45	0.35	0.73	-	2.2	-1.4	-2.11*
08C40	2012	yes	weak	27.8	-1.54	-0.93	-2.23	-	1.2	-.3	-2.08
08C58	2012	yes	weak	10	2.9**	0.64	5.05**	-	0	0	0.15
07C14	2012	yes	sizeable	9.5	4.04*	1.74	6.43**	-	1.3	.1	0.15

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ . .

Column  $T_d$  signals when the policy starts

Table 17: Estimates in the private sector for DRGs affected by the single tariff policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
12C08	2009	yes	no	96.2	2.62	1.95	3.81	3.03	6.5	4.5	1.22
03C15	2009	yes	no	94.8	0.77	0.4	1.69	0.69	4.5	3.5	0.71
03C21	2009	yes	no	40.2	-2.27	-0.78	-2.68	-4.1	1.4	-3	-2.56
03C09	2009	yes	no	15.4	-8.45**	-4.29**	-8.49**	-13.68**	1.1	-3	-2.67**
09C08	2009	yes	no	22.9	0.25	1.17	0.26	-0.56	.3	-3	-0.91
13C17	2009	yes	no	16.8	0.83	0.9	0.51	1.08	.1	0	1.3
08C28	2009	yes	no	37	1.56	0.65	2.15	1.02	.3	-1.6	0.12
09C13	2009	yes	no	79.9	-0.48	-0.69	-1.07	0.07	.5	0	1.5
13C08	2009 <sup>γ</sup>	yes	no	59.7	5.79*	2.6	7.34**	8.04*	0	.1	0.18
12C05	2009 <sup>α</sup>	yes	weak	5.8	2.12	0.02	1.53	4.77*	.1	0	-0.07
08C44	2009	yes	weak	83.4	2.34*	1.3	2.99*	3.19*	0	.1	0.52
13C20	2009	yes	weak	92	1.46	0.01	2.65	2.56	4.8	7.6	-0.63
12C13	2009	yes	sizeable	97.4	3.31	3.04	4.46	3.46	6.5	4.5	-0.13
13C12	2009	yes	sizeable	77.9	7.41*	3.59	8.72*	11.02*	5.2	-.9	-0.27
01C14	2009	yes	sizeable	68.9	4.56	2.26	5.26	6.45	3.5	-1.3	-2.07
06C12	2009	yes	sizeable	15.9	15.49**	1.54	17.99**	28.15**	12.9	-4.6	-5.46
09C07	2009	yes	sizeable	46.3	10.33*	4.13	10.95**	17.39*	5.2	4.1	-0.15
10C12	2012	yes	no	.3	0.33**	0.12	0.56**	-	0	0	0.03
10C10	2012	yes	no	2.8	0.15	-0.48	0.7	-	.3	-.2	0.2
02C06	2012	yes	no	76	3.57	0.63	6.84	-	.5	-1.3	2.95
10C07	2012	yes	no	3.8	-0.78	-0.58	-1.36	-	.2	-.1	-0.81
08C33	2012	yes	no	12.7	1.64	0.92	2.39	-	1.2	-.9	0.51
02C11	2012	yes	no	69.9	-8.5	-6.68*	-10.36	-	.8	-1.9	0.01
08C20	2012	yes	no	79	3.9	4.14	3.7	-	-.2	.2	-1.99
08C04	2012	yes	no	5.3	1.96	0.69	3.08	-	-1.2	-.1	0.98
13C04	2012 <sup>β</sup>	yes	no	3.2	0.13	-0.23	0.43	-	.1	0	-0.19
08C31	2012	yes	no	6	1.39	1.19	1.66	-	0	.1	0.03
21C06	2012	yes	weak	66.9	2.78	2.34	3.28	-	1.4	-1.2	-0.11
06C15	2012	yes	weak	6.2	0.16	0.16	0.34	-	.3	-.2	0.52
08C14	2012	yes	weak	75.6	1.16	0.59	1.53	-	.7	-.5	-1.56
08C58	2012	yes	weak	8.8	2.38**	0.35	4.43**	-	.4	-.2	-0.23

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ . .

Column  $T_d$  signals when the policy starts



Table 18: Estimates in the public sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo effect in $T_d - 1$
08C54	2008	no	weak	9.4	2.01	1.38	2.77	1.9	.3	-.2	-1.63
03C17	2008	no	weak	42.7	-1.67	-1.71	-2.02	-1.93	1.9	1.4	-4.47
08C38	2008	no	sizeable	44.7	9.49	5.38	10.49*	13.55*	4.8	3.3	-0.39
02C05	2008	no	sizeable	65.3	13.16*	6.95	17.32**	17*	8.2	7.3	2.5
08C45	2008	no	sizeable	60.1	11.51*	5.08	14**	16.67**	10.5	10.4	-0.07
06C19	2009	no	weak	8	2.28**	0.49	2.45**	4.1**	.1	0	0.69*
06C10	2009	no	weak	52.2	5.94**	5.25**	7.65**	6.15	-.3	.5	1.25
08C60	2009	no	weak	29	2.74	1.89	2.08	4.99	.9	-.8	-0.33
02C08	2009	no	weak	61.7	1.55	2.12	2.1	1.07	.2	.4	0.34
08C46	2009	no	weak	54.1	0.86	1.01	1.11	0.87	0	.4	1.01
06C14	2009	no	weak	12.7	2.34**	1.03**	2.12**	4.13**	.5	-.4	-0.35
13C19	2009	no	weak	11.7	1.85	-0.36	0.78	5.17	-.1	2	-0.15
12C07	2009	no	weak	17.5	0.99	0.35	0.79	2.15	.3	-.3	1.08*
13C11	2009	no	sizeable	64.8	5.43	3.03	4.41	9.08	2.4	1.9	1.27
13C10	2009	no	sizeable	25	11.14**	4.38	11**	19.11**	3.5	3.7	-0.51
13C05	2012	no	weak	1.5	0.66	0.53	0.79	-	-.1	0	0.16
13C09	2012	no	weak	18.4	2.92*	2.95**	3.15	-	-.5	.6	1.76**
11C07	2012	no	weak	25.8	0.2	-0.82	0.88	-	2	-1.6	-0.61

**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ . .

Column  $T_d$  signals when the policy starts

Table 19: Estimates in the private sector for DRGs only affected at first by the prior approval policy

DRG	$T_d$	Single tariff policy in $T_d$	Loose prior approval policy in $T_d$	Rate of same-day surgery in $T_d$	Mean effect	Effect in $T_d$	Effect in $T_d + 1$	Effect in $T_d + 2$	diff. in low potential rates	diff. in high potential rates	Placebo Effect in $T_d - 1$
08C34	2008	no	weak	.3	-0.18**	-0.06	-0.05	-0.44**	0	0	-0.02
03C06	2008	no	weak	38.4	0.41	1.53	-0.5	0.33	-.2	.3	4.94
08C54	2008	no	weak	16.1	1.17	0.89	1.44	1.49	.3	-.4	1.71
03C17	2008	no	weak	70.9	3.75	-0.22	4.86	6.55	1.9	-.6	-2.58
08C40	2009	no	weak	51.3	-2.42*	3.51	2.8	-13.03*	.1	-.1	0.43**
06C19	2009	no	weak	9.7	-1.14	-0.56	-0.92	-1.99	.4	-.3	-1.48
06C14	2009	no	weak	23.3	-1.19	-0.52	-1.37	-1.82	.9	-.7	-0.66
08C59	2009	no	weak	84.4	0.81	-0.07	1.8	1.43	3.6	-.6	-0.24
21C04	2009	no	weak	72	2.55	1.45	4.14	2.81	-.1	.1	-0.72
02C08	2009	no	weak	77.3	-0.48	-0.71	-0.33	-0.12	1.9	-.6	-3.24
08C46	2009	no	weak	76	1.25	0.31	1.89	2.09	0	.2	1.14
06C10	2009	no	weak	55.4	6.41*	-0.84	8.5*	11.48*	1.2	-.4	1.85
02C13	2009	no	weak	36.5	4.35	-1.09	3	10.91	11.2	-3.3	-3.24
13C19	2009	no	weak	7.1	3.26*	-0.36	2.89	7.39*	0	.3	-0.51
08C60	2009	no	weak	53.5	-0.14	-0.16	0.34	-0.46	1.2	-.9	-0.33*
12C07	2009	no	weak	27.4	4.73*	1.86	4.91*	7.68*	.1	-.2	1.97*
13C10	2009	no	sizeable	13.4	14.09**	2.21	15.88**	25.16**	7	0	-3.76*
13C05	2012	no	weak	1.3	-0.29	-0.51	-0.2	-	0	0	-0.24
09C14	2012	no	weak	18.6	3.25	1.71	5.46	-	.2	-.1	-1.6
08C35	2012	no	weak	5.9	0.08	-0.17	0.3	-	0	0	0.38
08C43	2012	no	weak	61.2	-0.02	-0.85	0.61	-	.2	-.2	0.13
03C07	2012	no	weak	27.7	0.97	0.25	1.71	-	3.1	-.8	-1.54
11C07	2012	no	weak	54.6	-0.3	0.56	-0.76	-	.3	-1.9	-1.3

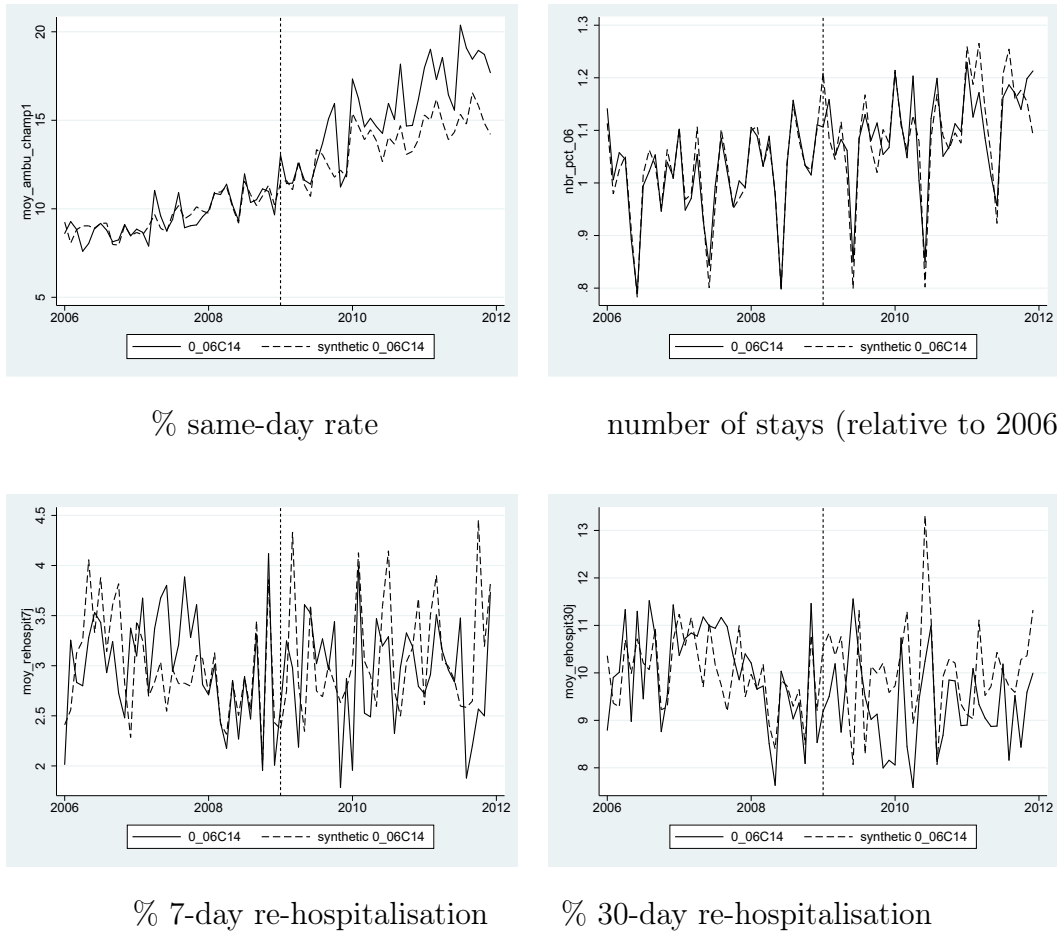
**Note:** Yearly estimates are the 12-month average estimates of the monthly difference between the incentivized DRG and its counterfactual DRG. They are expressed in percentage points.

**Lecture:** The significance thresholds are the following: \*  $p \leq 0.10$ , \*\*  $p \leq 0.05$ .

Column  $T_d$  signals when the policy starts

## F Graphs associated with some significant results regarding complementary variables

Figure 19: DRG 06C14 : Interventions for anogenital condyloma

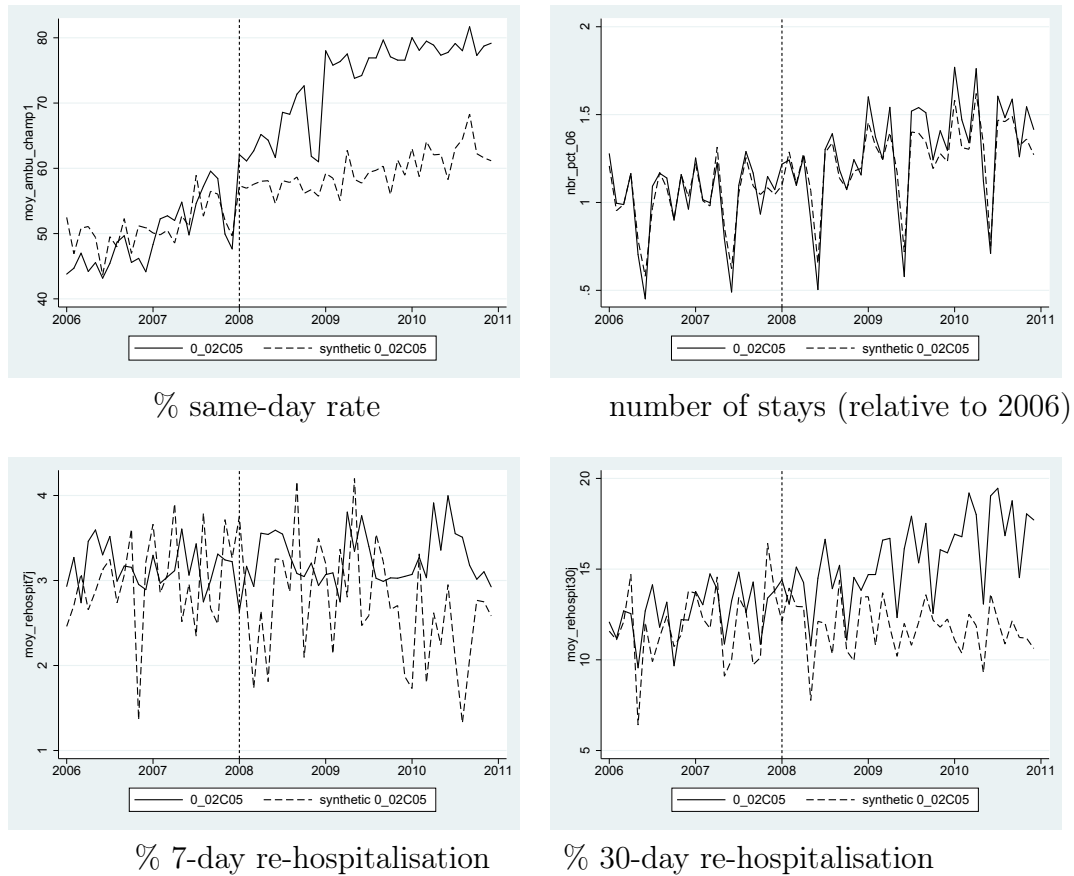


**Reading:** The incentivized DRG is in straight line and its synthetic control is in dashed line.

The vertical line in March 2009 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.

Figure 20: DRG 02C05 : Interventions on the lens with or without vitrectomy



**Reading:** The incentivized DRG is in straight line and its synthetic control is in dashed line. The vertical line in March 2009 stands for the date when the policy applied.

**Source:** ATIH, PMSI-MCO, treatments by the authors.

## **G The higher the distance to the mid-term prospect, the higher the effect: suggestive evidence.**

On average across all DRGs for which mid-term same-day prospects are available and which benefited from either policy in 2008/2009, the difference between the low bound of same-day prospects and the same-day surgery rate of these DRGs in the year when the policy was applied was of 19 % in the public sector, and of only 6% in the private sector (these figures rise respectively to 29% and 14% if rates of same-day surgery are compared with the high bound of same-day prospects). This is rough suggestive evidence that there are larger opportunities for improvements in same-day surgery rates in the public sector than in the private sector. It cannot be considered as more than suggestive evidence as the mid-term prospects that we use are the same for the public and the private sector, whereas patients ability to benefit from same-day surgery is likely to differ between sectors.

To check whether the larger the opportunity for improvement for same-day surgery, the higher the expected effect of either policy, we have regressed within each sector<sup>54</sup> the estimated (average) effect over the distance to the mid-term prospect (see Figure 21). The correlation is positive in both sectors. And it is higher when restricted to the DRGs for which the estimated effect was significant.

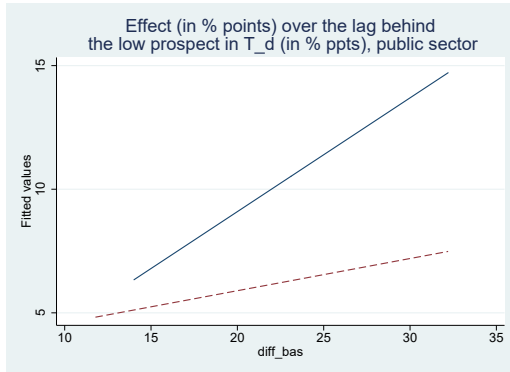
Note that when one considers the individual points,<sup>55</sup> this relation is not systematic. Notably, in both sectors, we occasionally do not find large effects for some DRGs for which much more same-day surgery could be achieved.

---

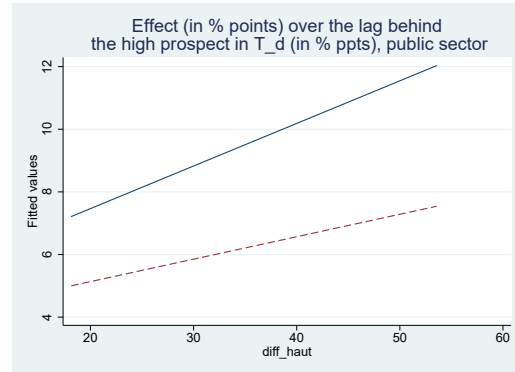
<sup>54</sup>We have taken into account each incentivized DRG for which a proper synthetic control was found.

<sup>55</sup>They are not reported, as, in order to get access to the prospects of same-day surgery, we committed not to disclose them and this would constitute a disclosure.

Figure 21: Regression of the estimated (average) effect over...

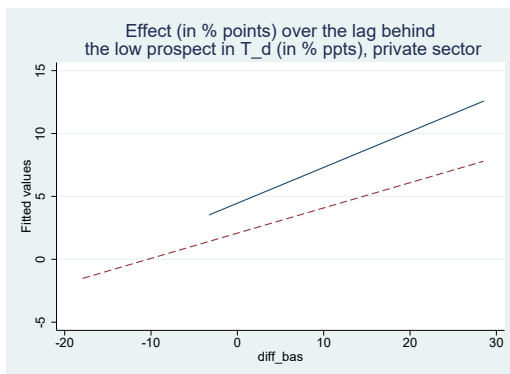


...the corresponding lower prospect

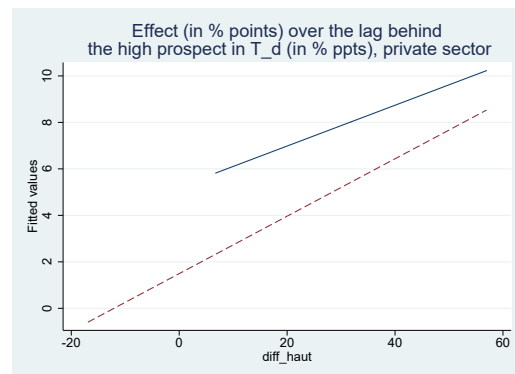


...the corresponding higher prospect

Public sector



...the corresponding lower prospect



...the corresponding higher prospect

Private sector

**Reading:** The dashed red line stands for the best linear fit between the average effect of each incentivized DRG for which a proper synthetic control was found, and the difference between its same-day prospect and its rate of same-day surgery on the first year of the policy. The solid blue line stands for the same best linear fit when restricted to the DRGs for which a robust significant effect has been found.

**Source:** ATIH, PMSI-MCO, medical learned societies, treatments by the authors.



## Chapitre 3

# Temps de transport domicile-travail, absentéisme et présentéisme en France



**Résumé :**

J'étudie les effets possibles du temps de trajet domicile-travail sur l'absentéisme et le présentéisme. En utilisant des modèles individuels à effet fixe sur des données de panel françaises récentes, et en tenant compte des changements d'emploi ou de lieu de résidence, ainsi que d'autres changements observables, je constate que le temps de trajet domicile-travail n'a pas d'effet significatif sur le présentéisme ou les arrêts-maladie en France. Ces résultats non significatifs dans de nombreux tests de robustesse. Je montre cependant que les deux relations devraient exister, au moins par l'intermédiaire de l'état de santé auto-évalué et du bien-être mental, mais que les grandeurs correspondantes sont probablement trop faibles pour être détectées. En outre, je montre que la stratégie consistant à restreindre l'échantillon aux personnes qui n'ont changé ni d'emploi ni de lieu de résidence, stratégie employée dans la littérature, ne résout pas nécessairement tous les problèmes d'endogénéité.

\* \* \*

---

# Does commuting time affect absenteeism and presenteeism in France? <sup>a</sup>

Alexandre Cazenave-Lacroutz <sup>b, c, d</sup>

## Abstract

*I investigate the possible effects of commuting time on absenteeism and presenteeism. Using individual fixed-effect models on novel French panel data, and taking into account changes of jobs or place of residence, as well as other observable changes, I find that commuting time has no significant effect on either presenteeism or sick leave absences in France. These non-significant findings persist across a large number of robustness tests. I do show, however that both relationships should exist, at least through the intermediation of self-assessed health and psychological well-being for which I find significant effects for women, but that the corresponding magnitudes are likely too small to be detected. In addition, I show that the strategy of restricting the sample to people who changed neither their job nor their place of residence employed in the previous literature does not necessarily solve the problem of endogeneity as claimed.*

**JEL:** I10, J22, R41.

**Keywords:** commuting time; absenteeism; sick leave; presenteeism.

---

<sup>a</sup>I thank Didier Blanchet, Inan Ceren, Eve Caroli, Brigitte Dormont, Andrew Clark, Yves Le Yaouanq, Tomáš Jagelka, Amelie Mauroux, Emilie Sartre, Joachim Winter for their suggestions and useful comments.

<sup>b</sup>Crest, 5 avenue Henry Le Chatelier, 91764 Palaiseau, France

<sup>c</sup>Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris

<sup>d</sup>Corresponding author: alexandre.cazenave-lacroutz@dauphine.eu

## 1 Introduction

Commuting is an important part of daily life for Western workers: in the US as well as in France for instance, they commute on average around 45 minutes per day.<sup>1</sup> These figures are rising (Goerke and Lorenz (2017), Coudène and Levy (2016)). Moreover, several studies in recent years have established its links with lower life satisfaction (Stutzer and Frey, 2008) or lower health satisfaction (Künn-Nelen, 2016),<sup>2</sup> with the psychological health of women (Roberts, Hodgson and Dolan, 2011; Munford et al., 2018), with absenteeism (Van Ommeren and Gutiérrez-i Puigarnau, 2011; Goerke and Lorenz, 2017; Künn-Nelen, 2016; Gimenez-Nadal, Molina and Velilla, 2018), and even with domestic violence (Beland and Brent, 2018). Far from being an object of interest only for the economics of transportation and the environment, commuting is part of the health and work economic nexus.

This paper provides new insight into the relationship between commuting and health at work, specifically absenteeism and presenteeism.

Absenteeism (being on sick leave) is a significant economic phenomenon which is on the rise. Though much has already been written about it (Johansson and Palme, 2005), it is still far from being clearly understood (Cazenave-Lacroutz and Godzinski, 2017). In particular, only four papers have recently studied whether commuting might cause more absenteeism, respectively in Germany (Van Ommeren and Gutiérrez-i Puigarnau, 2011; Goerke and Lorenz, 2017), Great Britain (Künn-Nelen, 2016) and the US (Gimenez-Nadal, Molina and Velilla, 2018). Various channels could induce such a relationship. First, longer commuting time could have adverse consequences on health (e.g. through more stress) and hence cause

---

<sup>1</sup>Using data of the Panel Study of Income Dynamics (PSID), Gimenez-Nadal, Molina and Velilla (2018) found that the average commuting time (over all workers, including those with no commute) was 43 minutes (resp. 38 minutes) for US male workers (resp. US female workers) over 2011-2015 ; in the data at my disposal, I found that the outward trip took only 23 minutes on average over 2013-2016 for a subset of French active workers.

<sup>2</sup>Note, however, that Dickerson, Hole and Munford (2014) and Lorenz (2018) did not find any impact of commuting on subjective well-being in general. However, Lorenz (2018) did find a (negative) impact on satisfaction with family life and leisure time.

more sick leave. Second, absenteeism could be understood (to some extent) as a decision to provide less labor.<sup>3</sup> As more commuting time decreases the time the employee can devote to work and leisure, and might also induce monetary costs,<sup>4</sup> it would reduce the financial return on a day's work, and thus the real daily wage. As such, it may have ambiguous effects on the choice between taking sick leave or going to work. But the modeling of these notions is quite diverse and could entail different results.<sup>5</sup> Although it does not deliver a definitive prediction for this empirical study, this theoretical corpus implies that there are also economic rationales such that commuting time might affect sick leave. It can be noted that all four empirical articles cited above found that more commuting increased absenteeism to some extent. Some of their results are contradictory, however, notably regarding gender differences or the magnitude of the effects.<sup>6</sup>

Conversely, the study of presenteeism, defined as the tendency of workers to attend employment despite illness (Pouliakas and Theodossiou, 2013; Chatterji and Tilley, 2002), is still in its infancy. Some evidence suggests that its economic costs can sometimes be even higher than the costs of absenteeism (Goetzel et al., 2004), and that it may trigger future absenteeism (Pichler and Ziebarth, 2017; Bergström et al., 2009). To the best of my knowledge, no previous paper has studied the relationship between presenteeism and commuting, probably due to data limitations.

---

<sup>3</sup>Such a choice might depend on the ability of the employee to go to see the doctor (or not) when sick, or to influence the doctor somewhat at the consultation, etc.

<sup>4</sup>These monetary costs might be caused by the gas bill for instance.

<sup>5</sup>Van Ommeren and Gutiérrez-i Puigarnau (2011) for instance reminds that "*When leisure time and work effort are substitutes, then the relationship between [absences] and the length of the commute is positive: see [Ross and Zenou (2008), proposition 1]. When they are complements, the relationship is negative.*" Closer to the provided intuition on labour supply, Gutiérrez-i Puigarnau and van Ommeren (2010) studies a model where workers can choose both the number of hours worked per day and the number of workdays. When workers face time and monetary costs that are both proportional to distance, they find that the effect of distance on number of workdays as well as total labour supply is ambiguous due to an increase in monetary costs that has both an income and a substitution effect. However, when modeling the effects of commuting time and monetary costs as independent exogenous variables, "*comparative static analysis entails then that an increase in commuting time increases daily labour supply, decreases the number of workdays, whereas it has an ambiguous effect on total labour supply.*"

<sup>6</sup>In Germany, Van Ommeren and Gutiérrez-i Puigarnau (2011) found that absenteeism would be about 15 to 20% lower if all workers had a negligible commute. In the US, however, Gimenez-Nadal, Molina and Velilla (2018) found that a 1% increase in the daily commute of male workers was associated with an increase of around 0.018% in sick-day absences per year. With an average daily commute of 43 minutes for men (back and forth), this implies that a 10-minute increase in commuting time would cause a 0.4% increase in sick leave.

This was explicitly acknowledged by Goerke and Lorenz (2017), who also pointed out that it would be worth addressing in future research. The direction of any such relationship is also ambiguous. As they found that long-distance commuting was associated with more absenteeism, Goerke and Lorenz (2017) expected "*long-distance commuters (...) to exhibit lower levels of presenteeism*". Yet such expectations that commuting time would have opposite effects on absenteeism and on presenteeism cannot be taken for granted: for instance, since more commuting might have adverse health effects, it could increase both absenteeism and presenteeism.<sup>7</sup>

I study these plausible relationships on a national sample of the French active population: the 2013 and 2016 waves of the French Working Condition Survey.<sup>8</sup> To control for unobserved individual heterogeneity, I exploit the panel dimension of the survey by including individual fixed effects. Following the literature, but with the advantage of a much more detailed questionnaire, I investigate whether taking potential time-varying cofounders into account might influence the results.

In contrast to previous research, I do not find that commuting leads to a significant increase in sick leave and, contrary to the expectations of Goerke and Lorenz (2017), I do not find that commuting time decreases presenteeism. Rather, the point estimates are almost always positive (and non-significant). These two (non-significant) findings are robust to alternative specifications, improvements made to the identification strategy, or focus on the largest changes. In addition, no robust heterogeneity is found regarding sector, gender or any other investigated variables. I find that commuting time significantly decreases psychological well-being, especially of women (in line with existing research), and also their self-assessed health

---

<sup>7</sup>Even more generally, Arnold, de Pinto et al. (2015) also noted that work-related characteristics can have a distinct or similar impact on presenteeism and on absenteeism. To explain it, they resorted to another argument than health effects: they hypothesized that work-related characteristics may affect the worker's absence decision but also the individual-specific sickness definition.

<sup>8</sup>In 2013: *Enquête Conditions de travail*; in 2016, its full name is: *Enquête Conditions de travail et Risques psychosociaux*.

(though less significantly so). Both are also found to significantly increase both absenteeism and presenteeism. The corresponding effects are small in magnitude, however. Both studied relationships should therefore exist, at least through the intermediation of self-assessed health and psychological well-being, but the corresponding magnitudes are likely too small to be detected. From a methodological viewpoint, I also show that the state-of-the-art strategy in this field of research - considering within-changes in commuting involving neither a change of job nor of place of residence - does not cause variations in commuting to be driven only by exogenous moves of the employer.

Moreover my preferred estimates point to a 1-minute elasticity of absenteeism of only 0.001, and a 1-minute elasticity of presenteeism of 0.002, implying that one standard deviation in commuting time (20 minutes, which is also close to the mean commuting time) would translate into only 2% more sick leave and 4% more presenteeism. This assertion is somewhat confirmed by my examination of the two direct health variables at my disposal: self-assessed health and psychological well-being. For both variables, I find that commuting time would have a small detrimental health effect, specifically on women. Hence, if health effects are the main causal channel between commuting time and absenteeism/presenteeism, one should not expect large effects on these variables, as psychological well-being and self-assessed health are not greatly affected.

This study has three main advantages over previous papers studying the relationship between commuting and health at work. First, no previous researcher has been previously able to study the possible impact of commuting time on presenteeism, although it has been explicitly cited as an area where research on commuting is needed. Second, I can study presenteeism jointly with absenteeism, which can improve our understanding of these two related variables. Third, in addition to using the conventional fixed-effect strategy, I am able to exploit the richness of the French Working Condition Survey questionnaire to understand what might

explain part of the changes in commuting time. Finally, this is the first research paper using French data on the commuting-absences relationship.

The remainder of the paper is organized as follows. Section 2 presents the data. Section 3 describes the empirical strategy. Section 4 presents the results. Section 5 focuses on the possible intermediation of two health variables, which is to say self-assessed health and psychological well-being. Section 6 concludes.

## **2 Data and variables**

I use the 2013 and 2016 waves of the French Working Condition Survey. The 2013 sample was initially drawn to get a representative sample of the active population, with additional individuals being sampled in the public sector and in the hospital sector. 34,000 individuals were interviewed between October 2012 and February 2013. Almost all the individuals who were interviewed were contacted again for the second wave.<sup>9</sup> 20,000 (still active) individuals were eventually re-interviewed between October 2015 and June 2016.<sup>10</sup> To my knowledge, it is the only nationwide panel survey for France to include information relative to commuting, as well as information on presenteeism and sick leave absences. It also includes rich information on labor market status, health, family life and socio-economic variables.

The French Working Condition Survey provides self-reported measures of the annual number of days of sick leave (not including maternity leave) and of the annual number of times<sup>11</sup> the interviewed person went to work thinking she should

---

<sup>9</sup>For budget reasons, only a third of the 4,800 interviewees living overseas in 2013 were contacted again in the 2016 wave.

<sup>10</sup>The private non-hospital sector was underrepresented in this final sample. There was also attrition between the two waves. For the descriptive statistics, I use weightings from the survey producer (the Statistical Department of the Labor Ministry) to make the sample representative of the 2013 active population that was still active in 2016.

<sup>11</sup>Interviewer are told that "once" was one day.

have stayed home due to sickness (Inan, 2020).<sup>12</sup> On average, people have been absent for sick leave for almost 8 days in the twelve months preceding the interview (see Table 1). Over the same period, they have gone to work despite being sick for almost 3 days. Both averages hide considerable heterogeneity, as more than half of the respondents did not report any sick leave absence or presenteeism. Though these raw statistics do not control for job characteristics, women reported both more sick leave absences and more presenteeism than men.

**TABLE 1.**  
Sick leave absences and presenteeism: descriptive statistics

variable Y	mean E(Y) (in days)	P(Y > 0)	p(75) (in days)	p(90) (in days)	p(95) (in days)	conditionnal mean E(Y>0) (in days)
All						
absences $A_i$	7.9	32 %	3	15	40	24.9
presenteeism $P_i$	2.8	42 %	3	5	10	6.5
Men						
absences $A_i$	6.6	29 %	2	14	30	23.0
presenteeism $P_i$	2.0	36 %	2	5	10	5.7
Women						
absences $A_i$	9.2	35 %	4	21	45	26.5
presenteeism $P_i$	3.5	49 %	3	8	10	7.1

*Source:* French Working Condition Survey 2013-2016, 18-65 year old wage earners : in both 2013 and 2016. Weights are used.

*Lecture:* p(x) is the  $x^{th}$  percentile; P(Y > 0) is the probability that Y > 0.

Each respondent was also asked how long their commute took.<sup>13</sup> The interviewers were told this concerned their outward travel from home to the main workplace, and without taking any detours into account.<sup>14</sup> As shown in Table 2, respondents spent an average of 24 minutes on their outward commute over the period. This is a little higher for men than for women. Although 70% to 80% of

<sup>12</sup>The wording differed between respondents working at home - "*In the last 12 months, have you ever worked while thinking that you shouldn't have because you were sick?*" - and the other - "*In the last 12 months, have you ever gone to work thinking that you should have stayed home because you were sick?*".

<sup>13</sup>The question was: "*On average, how long does it take you to get to work?*"

<sup>14</sup>If there was no usual trip, the interviewer was to note time travel as "997". I discarded these observations.



the sample report almost no change between the two waves,<sup>15</sup> more than 15% of respondents reports a change of more than 15 minutes.

**TABLE 2.**  
Commuting time: levels and within-individual changes

	all	men	women
Commuting Time $T_i$ (in min.)			
mean	24.3	24.9	23.6
sdt. dev.	20.1	20.5	19.7
p(5)	3.0	3.0	3.0
p(10)	5.0	5.0	5.0
p(50)	20.0	20.0	20.0
p(90)	45.0	50.0	45.0
p(95)	60.0	60.0	60.0
Change in Commuting Time $\Delta T_i$ (in min.)			
mean	0.5	0.8	0.2
sdt. dev.	16.5	17.0	16.1
p(5)	-20.0	-20.0	-20.0
p(10)	-10.0	-10.0	-10.0
p(50)	0.0	0.0	0.0
p(90)	15.0	15.0	15.0
p(95)	25.0	25.0	23.0
% with exactly no change	38.2 %	37.4 %	39.0 %
% with change no higher than 5 minutes	69.5 %	69.0 %	69.9 %
% with change no higher than 10 minutes	79.6 %	79.5 %	79.7 %
% with change no higher than 20 minutes	89.9 %	89.5 %	90.2 %
% with change no higher than 30 minutes	94.4 %	93.8 %	95.0 %

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for which the focal variable (commuting time) and the core controls are available. Weights are used.

*Lecture:* p(x) is the  $x^{th}$  percentile ; std. dev. is the standard deviation.

Following the literature, I consider a large number of core explanatory variables. Some are related to personal characteristics such as gender and region dummies, the fact of living in a couple, having children, having children under the age of three, age, and educational attainment. I also incorporate core job-related variables, such as tenure and type of contract, occupational position, firm size and industry dummies, belonging to the public sector, and belonging to the hospital sector. A precise description of the controls, along with some descriptive statistics can be found in Annex A.

<sup>15</sup>Due to noisy answers, I do not consider that true absence of change concerns only those with exactly no change between the two waves, but rather those whose reported change is less than 5 minutes or less than 10 minutes. All three statistics are however reported in Table 2.

In addition to these core controls, I consider four additional controls that may be also affected by commuting time: working time and (log of) income, self-assessed health status, and psychological well-being. Although controlling the dependent variables by variables that may also be affected by commuting time might bias the total effect of commuting time, the literature dealing with the health consequences of commuting is notably "*inconsistent*" in terms of including potential compensating factors of commuting such as income and working time (Künn-Nelen, 2016).<sup>16</sup> Similarly, health status is sometimes included in the explanatory variables (Gimenez-Nadal, Molina and Velilla (2018), Goerke and Lorenz (2017)), sometimes not (Van Ommeren and Gutiérrez-i Puigarnau (2011)) and is even sometimes another dependent variable (Künn-Nelen, 2016). I choose not to include them in my main specification, but to control for them in additional regressions. As some results also point to an impact of commuting on psychological health (Roberts, Hodgson and Dolan (2011), Dickerson, Hole and Munford (2014)) and the classic measure of psychological well-being defined by the World Health Organization (WHO) (Topp et al., 2015) are available in the two waves of the survey, I also consider it among these additional controls. It is a 25-point Likert scale depicting subjective psychological well-being based on 5 questions regarding how often the interviewed person has had certain feelings in the two weeks before the interview.<sup>17</sup> The higher the score, the higher the psychological well-being. Similarly, the higher the health status, the better.

Finally, I restrict my main sample to 18-65 year old wage earners for whom

---

<sup>16</sup>They are dubbed "compensating factors" because an employee may have accepted to do a job that imposes a long commute notably because this job brings her a higher wage or shorter working time compared to alternative offers. These issues are a focus of the literature concerned by the so-called "commuter paradox" (See Stutzer and Frey (2008) ; or footnote 35.), whereas I am primarily interested in the existence of a net impact of commuting on absenteeism and presenteeism.

<sup>17</sup>The 5 feelings are the following: "*I felt good and in a good mood*"; "*I felt calm and quiet*"; "*I felt full of energy and vigorous*"; "*I woke up feeling fresh and refreshed*"; "*My daily life has been filled with interesting things*". Answer to each questions could be: 1. *All the time*, 2. *Most of the time*, 3. *More than half of the time*, 4. *Less than half of the time*, 5. *Occasionally*, 6. *Never*, 9. *Don't Know*.

the focal variable (commuting time) and core controls are available. This notably exclude self-employed workers, as do all the articles in the literature for fear of their sick leave differing from that of employees. I also exclude the very few respondents with more than three hours of commuting time,<sup>18</sup> as well as those with more than 3 months of sick leave absence in either wave.<sup>19</sup> This yields a balanced panel of 15,519 people observed twice.

### 3 Empirical strategy

In most of the paper, I adopt a fixed-effect strategy to assess the effect of commuting time on absences and presenteeism empirically. This controls for any time-invariant idiosyncratic effects that may confound the relationship under study.<sup>20</sup>

My main specification has the following general form:

$$Y_{i,t} = \alpha.T_{i,t} + \beta.x_{i,t} + \mu_i + \nu_{2016} + \epsilon_{i,t} \quad (1)$$

The dependent variable  $Y_{i,t}$  is the annual number of days of sick leave or, respectively, of presenteeism of employee  $i$  as reported in wave  $t$ . As both dependent variables are count variables, I mainly use the Poisson distribution since "*the Poisson distribution is the nominal distribution for count data in much the same way that the normal distribution is the nominal distribution for unbounded, continuously distributed data*" (Wooldridge, 1999a). As my dependent variables are

---

<sup>18</sup>That is: probably more than 6 hours in a day. This concerns 20 observations, one of which has an outward commute of more than 9 hours.

<sup>19</sup>This concerns 270 observations. These outliers have presumably endured a significant health shock, that may also have had consequences on their commuting time this year, which would be reverse causality. In case of a significant finding regarding presenteeism, a similar restriction has to be made as a robustness test.

<sup>20</sup>Such bias was observed by Künn-Nelen (2016), for instance, who did not find significant effects of commuting on health in OLS regressions or in errors-in-variable regressions, while she did find significant effects in fixed-effects specifications.

over-dispersed,<sup>21</sup> I also consider the Negative Binomial distribution as an alternative specification. In technical Annex B, I outline why the Poisson distribution remains my preferred specification.<sup>22</sup> For both, I use the log-function as the link function (see Annex B).

$T_{i,t}$  is the interviewee's commuting time when interviewed in wave  $t$ .  $\alpha$  is the coefficient of interest.  $x_{i,t}$  stands for the set of controls defined in Part 2 (see also Annex A), eventually supplemented by new controls (that will be described below).  $\mu_i$  is an individual (employee) fixed effect. It controls for unobserved time-invariant individual-specific heterogeneity.  $\nu_{2016}$  is a fixed-effect for the 2016 wave.  $\epsilon_{i,t}$  is a heteroskedastic error term, whose (robust) variance is also clustered at the employee level.

Such a fixed-effect strategy controls for time-invariant unobserved characteristics, but not for time-varying unobserved characteristics correlated with both the dependent and explanatory variables. Possible endogenous changes in both commuting time and absences/presenteeism are thus an object of concern in the literature, and some authors have tried to fix such concerns by using only observations with the same job and the same household location during the period of observation.<sup>23</sup> Following Gimenez-Nadal, Molina and Velilla (2018), I prefer to add dummies equal to 1 in 2016 when there was a change of place of residence or of job. This should capture the average unobserved changes in job quality (for instance ) that might be associated with a job change. For the purpose of robustness however, I restrict the sample to those not having changed either job or place of residence some specifications of Section 4.3, rather than simply having them in controls.

<sup>21</sup>When applying pooled Negative Binomial regressions in Annex D, the alpha-constant was significantly higher than 1 for both variables, which is evidence of over-dispersion.

<sup>22</sup>In practice however, this is not very important as both specifications (the Poisson one and the Negative Binomial one) displayed similar results most of the time.

<sup>23</sup>According to Goerke and Lorenz (2017), "*A variation in commuting distance will, thus, only occur if a firm alters its location*". According to Van Ommeren and Gutiérrez-i Puigarnau (2011), this would make them use "*changes in commuting distance that are employer-induced, which guarantees that commuting distance is exogenous*".

Using the richness of the survey questionnaire, I also show in Annex G that, even with such restrictions, there is still ample room for various kinds of changes plausibly affecting both the dependent and focal variables (other than employer-induced firm relocation). More precisely, as depicted by summary Table 3, only 5% to 20% of large changes in the commuting time of the restricted sample might be attributed to an exogenous change in firm location; and I am able to control for plausible endogenous reasons for a large change in commuting time for 50% of the large changes in commuting time between the two waves. To further assess the robustness of the main specification relative to concerns of omitted variable bias, I complement it with a new set of variables to take time-varying factors into account as much as possible (see Table G.IV). In Section 4.3, I apply it both to the restricted sample and to the unrestricted sample.

Following Van Ommeren and Gutiérrez-i Puigarnau (2011), it should be noted that all the estimates might be downward biased due to the difference of time frame between my dependent variables (counts over the last 12 months) and my focal variable (commuting at the time of the interview). However, when Van Ommeren and Gutiérrez-i Puigarnau (2011) took this into account, it made no empirical difference. Moreover, as detailed in Annex F, if there is any such bias, it will arguably be much smaller in my case than it would have been for Van Ommeren and Gutiérrez-i Puigarnau (2011).

**TABLE 3.**  
Plausible reasons behind large changes in commuting time in the restricted sample

Question or subset of questions	% of $ \Delta T_i  \geq 15min$
<b>C1 - Relocation or restructuring of the firm</b>	<b>20 %</b>
<i>The address of the establishment where you work is [address provided in wave 1]. Is that correct? → No, the firm moved</i>	5.0%
<i>In the last twelve months, has your working environment been significantly affected by a restructuring or a relocation of the [firm] where you work? → Yes.</i>	17.7%
<b>C2 - In a job where the workplace varies often (in 2016 or in 2013)</b>	<b>25 %</b>
<b>C3 - The respondent seems not to have been truly living in the same home at both waves.</b>	<b>1 %</b>
<b>C4 - A possible change in job previously unnoticed.</b>	<b>17 %</b>
<b>all: C1 or C2 or C3 or C4</b>	<b>49%</b>

## 4 Results

### 4.1 Descriptive pooled regressions

Before presenting the results of the fixed-effect regressions, I begin by documenting the descriptive results from pooled regression- in Table 4 and 5. They show that when only the levels are considered to investigate the relationship of interest (that is: when time-unvarying unobserved characteristics are not controlled for by differentiating all variables), commuting time increases both sick leave and presenteeism. Yet, the estimated magnitudes are not very large, especially for the impact on sick leave.

In practice, Table 4 reports the association between commuting time and sick leave absences when using pooled Poisson specifications in columns (1), (2) and (3), or pooled Negative Binomial regressions in columns (4), (5) and (6). Columns (1) and (4) have only the classic core controls presented in Annex A (e.g. gender, living within a couple). The remaining regressions include additional controls that may themselves be affected by commuting. Working hours and (log of) wage are

included in columns (2) and (5). Self-reported health status and psychological well-being are furthermore added in columns (3) and (6). Only the coefficients of interests are reported.<sup>24</sup>

Due to their very low magnitude, the results are easy to interpret. For instance, the 0.00195 coefficient in column (1) of Table 4 implies that an increase by 10 minutes in commuting time causes an increase of  $0.00195 \times 10$  in the logarithm of the number of days on sick leave. That is, it causes an increase by  $\exp(0.0195)$  in the number of days in sick leave, i.e. a 2% increase.<sup>25</sup> Hence an increase of 20 minutes in commuting time (that is: by a standard deviation or by approximately a mean commuting time, as shown in Table 2), is associated with an increase of only 4% in sick leave (and of almost 10% in presenteeism as can be read from Table 5). The corresponding 95% confidence interval would represent an increase in sick leave of between 1.1% and 6.6% (respectively an increase in presenteeism of between 5.1% and 14.1%) for 20 additional minutes. The effects are somehow similar with the Negative binomial specification (see Column (4)).

In these descriptive regressions, commuting time would increase both absenteeism and presenteeism. This goes against the intuition of Goerke and Lorenz (2017) quoted in the Introduction. I see at least two explanations behind this positive correlation. First, commuting time could be correlated with (or could cause) bad health that may in turn cause both absenteeism and presenteeism. For instance, many studies report a correlation between car driving and low back pain (Gallais and Griffin, 2006). The latter can be a major cause of sick leave (Ree et al., 2016) and arguably can also cause presenteeism. As commuting in France was mostly by car over that period (Bolusset and Rafraf, 2019), such a relationship is a channel by which more commuting time can be positively correlated with both sick leave and presenteeism. Second, work-related characteristics (such as

---

<sup>24</sup>Annex D also presents all the coefficients corresponding to columns (3) and (6).

<sup>25</sup>A first-order development of the exponential yields:  $\exp(0.0195) \simeq 1 + 0.0195$ . This is a reasonable approximation. For instance  $\exp(0.0195)$  is in fact close to 1.0197.

commuting time) could also affect the individual-specific sickness definition in addition to their absence decision (Arnold, de Pinto et al., 2015). For instance, even without negative consequences on health, the stress endured in traffic jams could both negatively affect the decision to come to work and the feeling of presenteeism in case of work.

These effects significantly<sup>26</sup> increased when adding income and the number of working hours to the regressions (see the difference between column (1) and column (2)). There was a similar increase with the Negative Binomial specification (see the difference between column (4) and column (5)). This indicates that these two additional variables correlate with commuting time and that they explain some variation in sick leave absences, as shown in columns (2) of Tables 4 and 5. For instance, a higher income is associated with less sick leave and less presenteeism. This could make the coefficients of interest increase between regressions if people with more commuting time also tend to have higher incomes. This is indeed the case, as shown by Table E.II of Annex E.

When further adding the self-assessed health and psychological well-being variables to the regressions, the coefficients of interest are significantly diminished. It even loses any significance in the case of absenteeism with the Poisson specification (see column (3) of Table 4). As commuting time does correlate negatively with both self-assessed health and psychological well-being (as shown by Table E.II of Annex E), and both self-assessed health and a higher psychological well-being decrease both absenteeism and presenteeism, it can be concluded that part of the correlation between commuting time and both absenteeism and presenteeism is driven by the lower psychological well-being and worse health of those with long commuting times.<sup>27</sup>

---

<sup>26</sup>See Annex C. According to a Hausman test, the 0.0005 increase in the Ti coefficient was significant at the 1% threshold.

<sup>27</sup>I investigate in Section 5.2 whether these latter relationships might themselves be causal, or if they reflect mere correlations.



**TABLE 4.**  
Descriptive pooled regressions on the impact of commuting time on absenteeism

Specification	Absenteeism $A_i$					
	(1) Poisson	(2) Poisson	(3) Poisson	(4) Negative Binomial	(5) Negative Binomial	(6) Negative Binomial
commuting time $T_i$	0.00195** (0.000902)	0.00245*** (0.000899)	0.00136 (0.000893)	0.00238** (0.000932)	0.00319*** (0.000947)	0.00230** (0.000994)
log(monthly income)		-0.256*** (0.0379)	-0.184*** (0.0402)		-0.372*** (0.0611)	-0.277*** (0.0603)
working hours		0.00132 (0.00269)	-0.00118 (0.00263)		0.00189 (0.00320)	-0.00110 (0.00332)
good health			-0.506*** (0.0218)			-0.490*** (0.0237)
psychological well-being			-0.0146*** (0.00341)			-0.0165*** (0.00384)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26584	26584	26584	26584	26584	26584
Fixed-effects	No	No	No	No	No	No

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (gender, region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. To make regressions fully comparable, they are also not included in the pooled regressions. Standard-errors are robust and clustered.

**TABLE 5.**  
Descriptive pooled regressions on the impact of commuting time on presenteeism

Specification	Presenteeism $P_i$					
	(1) Poisson	(2) Poisson	(3) Poisson	(4) Negative Binomial	(5) Negative Binomial	(6) Negative Binomial
commuting time $T_i$	0.00493*** (0.00158)	0.00501*** (0.00157)	0.00304** (0.00155)	0.00485*** (0.00128)	0.00515*** (0.00130)	0.00250** (0.00109)
log(monthly income)		-0.138** (0.0701)	-0.0226 (0.0726)		-0.131* (0.0682)	-0.0686 (0.0639)
working hours		0.0167*** (0.00424)	0.0109** (0.00439)		0.0173*** (0.00354)	0.0151*** (0.00391)
good health			-0.517*** (0.0418)			-0.469*** (0.0306)
psychological well-being			-0.0715*** (0.00514)			-0.0856*** (0.00416)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26242	26242	26242	26242	26242	26242
Fixed-effects	No	No	No	No	No	No

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (gender, region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. To make regressions fully comparable, they are also not included in the pooled regressions. Standard-errors are robust and clustered.

## 4.2 Simple fixed-effect regressions

I now switch to the fixed-effect specification to control for time-unvarying unobserved characteristics and bring the estimates of interest closer to a causal identification of an effect of commuting time on absenteeism and presenteeism. All the coefficients of interest are lower in the fixed-effect specifications (in Tables 6 and 7) than in the pooled regressions (in Tables 4 and 5), both for absenteeism and for presenteeism. It can be noted that they are not significantly lower for absenteeism.<sup>28</sup> Regarding the impact of commuting time on absenteeism, rather than causing a large fall, the incorporation of the fixed-effects mostly increases the standard errors considerably and/or causes a fall in the coefficients of interest, yielding insignificant results in 5 specifications out of 6. Similarly, all the results associated with presenteeism are also insignificant with the fixed-effects. When the Negative Binomial specification is used, the loss of significance is also due to significant falls in the coefficients.

To provide a better understanding of these results against the literature finding that commuting time may have sizeable effects on sick leave, I consider the 95% confidence interval associated with the coefficient of interest in column (1). It is  $[-0.0036; 0.0042]$  for absenteeism (respectively  $[-0.0036; 0.0071]$  for presenteeism). It implies that at the 5% threshold, we can reject the hypothesis that a 20-minute increase in commuting time increases sick leave by more than 9% (respectively, increases presenteeism by 13%). Hence, although the estimates are very small in magnitude, I cannot completely rule out that commuting time may have sizeable effects on either sick leave or absenteeism.

Adding supplementary explanatory variables may have different effects according to the specifications. When adding the (log of) income and the number of working hours to the regressions with absenteeism, the coefficient of interest increases

---

<sup>28</sup>See Annex C. Regarding absenteeism, the 0.001 decrease in the  $T_i$  coefficient between the first columns of the two tables is for instance not significant at the 10% threshold. Regarding presenteeism, the 0.003 decrease in the  $T_i$  coefficient between the first columns of the two tables is significant at the 10% threshold only.

a bit for the Poisson specification (from 0.000280 to 0.000523) but decreases a bit for the Negative Binomial Specification (from 0.00219 to 0.00217). This might notably come from the absence of clear independent impact of income on absenteeism and presenteeism (the sign of the corresponding coefficient varies between regressions and only one is significant - at the 10% threshold)

When further adding self-assessed health and psychological well-being, the coefficient of interest significantly decreases in the Negative Binomial Regression for absenteeism (from 0.00217 to 0.00144), as previously observed in Table 4 but it (non-significantly) increases in the Poisson regression for absenteeism (from 0.000523 to 0.000615). For presenteeism, it still decreases in both specification (the decrease is significant only in the Negative Binomial specification). These results hint at a possible intermediation by the health status and the psychological well-being of the (eventual) impact of commuting time on absenteeism and on presenteeism.

**TABLE 6.**  
Use of fixed-effect models to study the impact of commuting time on absenteeism

Specification	Absenteeism $A_i$					
	(1) Fixed-Effect Poisson	(2) Fixed-Effect Poisson	(3) Fixed-Effect Poisson	(4) Fixed-Effect Negative Binomial	(5) Fixed-Effect Negative Binomial	(6) Fixed-Effect Negative Binomial
commuting time $T_i$	0.000280 (0.00199)	0.000523 (0.00201)	0.000615 (0.00198)	0.00219 (0.00151)	0.00217** (0.00108)	0.00144 (0.00106)
log(monthly income)		-0.186* (0.105)	-0.150 (0.107)		0.00393 (0.0646)	0.0413 (0.0639)
workin hours		0.0112** (0.00552)	0.0113** (0.00541)		0.00574* (0.00316)	0.00423 (0.00276)
good health			-0.327*** (0.0363)			-0.307*** (0.0273)
psychological well-being			-0.0224*** (0.00593)			-0.0236*** (0.00418)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12392	12392	12392	12392	12392	12392
Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered. For the Negative Binomial Regressions (NegBin), the clustered robust standard errors are estimated through 50 bootstrap replications.

**TABLE 7.**  
Use of fixed-effect models to study the impact of commuting time on  
presenteeism

Specification	Presenteeism $P_i$					
	(1) Fixed-Effect Poisson	(2) Fixed-Effect Poisson	(3) Fixed-Effect Poisson	(4) Fixed-Effect Negative Binomial	(5) Fixed-Effect Negative Binomial	(6) Fixed-Effect Negative Binomial
commuting time $T_i$	0.00175 (0.00271)	0.00207 (0.00274)	0.000821 (0.00235)	0.000913 (0.000749)	0.000890 (0.000916)	0.000134 (0.00104)
$\log(\text{monthly income})$		0.0630 (0.0939)	0.0791 (0.0998)		-0.0155 (0.0535)	0.0327 (0.0629)
working hours		0.0174** (0.00682)	0.0136** (0.00642)		0.0124*** (0.00417)	0.0107*** (0.00301)
good health			-0.377*** (0.0462)			-0.200*** (0.0216)
psychological well-being			-0.0439*** (0.0100)			-0.0465*** (0.00326)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15334	15334	15334	15334	15334	15334
Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered. For the Negative Binomial Regressions (NegBin), the clustered robust standard errors are estimated through 50 bootstrap replications.

### 4.3 Results when identification is improved

Even though fixed-effect specifications prevent bias coming from time-unvarying characteristics, they do not prevent bias coming from misleading inference due to time-varying unobserved characteristics or changes. If commuting time increases sick leave absences, changes in job or in place of residence might bias the estimated coefficients downwards. For instance, increases in commuting time might be associated with better jobs (which is unobserved), moderating the effect of commute time on the dependent variables.

Changes in the place of residence between the two waves are recorded and one question in second wave asked for any change of job (more precisely, a change of profession or of employer). I also have access to the starting year of the contract: as the first wave was conducted in 2013, a contract starting after 2013 might be indicative of a change of job between the two waves (see Annex G for all details). Similarly to Gimenez-Nadal, Molina and Velilla (2018), the results are not significantly altered when I take these changes into account:<sup>29</sup> neither for sick leave absences (in column (2) of Table 8), nor for presenteeism (in column (2) of Table 9).

In columns (3) of these tables, I restrict the sample to the respondents who has changed neither place of residence nor job between the two waves, rather than controlling for it. This decreases the coefficients and increases standard errors (perhaps due to fewer observations). The results remains low and non-significant. This is also the case when adding in columns (4) new time-varying characteristics (or impulse variables) that provides a plausible explanation behind 50% of the large changes in commuting time in the restricted sample (see Annex G). In particular, of these 50%, only 5% (i.e. 10% of 50%) come for sure from a relocation of the firm (and at most 20%), although some previous authors though they would explain most of the changes. Other plausible explanations include working in a job where the workplace varies often, or unnoticed changes in job or place of residence.

---

<sup>29</sup>I create three dummies equal to 1 in wave 2 whenever there is such a change. They form the so-called "ChangeVariables".

In case the differences in standard errors between columns (2) and (3) are mainly due to the number of observations, I come back to the unrestricted sample in columns (5), with job changes and home changes being incorporated as dummy variables (as in column (2)). At that point, this is my estimate of reference. It may be noted that the corresponding coefficient is the highest, both for sick leave absences and for presenteeism, and that the corresponding standard error is also the lowest for both sick leave absences and presenteeism. Despite this, both coefficients are still low and non-significant.

Considering the 95% confidence interval, it is  $[-0.0024; 0.0047]$  for the impact on absenteeism, and  $[-0.0023; 0.0063]$  for the impact on presenteeism. At the 95% threshold, we can therefore be confident that a 20-minute increase in commuting time would yield at most an increase of 9% in sick leave, and of 12% in presenteeism.



**TABLE 8.**  
Main results for absenteeism

Specification	Absenteeism $A_i$				
	(1) Fixed-Effect Poisson	(2) Fixed-Effect Poisson	(3) Fixed-Effect Poisson	(4) Fixed-Effect Poisson	(5) Fixed-Effect Poisson
commuting time $T_i$	0.000829 (0.00184)	0.000966 (0.00184)	0.000513 (0.00218)	0.000592 (0.00217)	0.00112 (0.00182)
Controls	Yes	Yes	Yes	Yes	Yes
ChangeVariables	No	Yes	No	No	Yes
AdditChangeVariables	No	No	No	Yes	Yes
Observations	14206	14206	11858	11858	14206
Restricted sample	No	No	Yes	Yes	No

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. In the restricted sample, I keep only the respondents not changing their job or place of residence between waves.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. It is also the case of tenure and firm size dummies in the restricted sample. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

**TABLE 9.**  
Main results for presenteeism

Specification	Presenteeism $P_i$				
	(1) Fixed-Effect Poisson	(2) Fixed-Effect Poisson	(3) Fixed-Effect Poisson	(4) Fixed-Effect Poisson	(5) Fixed-Effect Poisson
commuting time Ti	0.00175 (0.00241)	0.00172 (0.00241)	0.00127 (0.00325)	0.00165 (0.00296)	0.00198 (0.00219)
Controls	Yes	Yes	Yes	Yes	Yes
ChangeVariables	No	Yes	No	No	Yes
AdditChangeVariables	No	No	No	Yes	Yes
Observations	17494	17494	14500	14500	17494
Restricted sample	No	No	Yes	Yes	No

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. In the restricted sample, I keep only the respondents not changing their job or place of residence between waves.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. It is also the case of tenure and firm size dummies in the restricted sample. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

## 4.4 Dealing with eventual data issues

Changes in the commuting time variable might be noisy. Indeed, commuting time in itself is probably noisy, and there is more measurement error in the regressors of a fixed-effects equation than in the levels of the regressors (see for instance Angrist and Pischke (2008), chapter 5). Hence, small changes might be just a reflection of measurement error, biasing downward the relation of interest. Hence, I also identified its potential effects for large changes only.

The threshold to define a change large enough not to be caused by measurement error is necessarily somewhat arbitrary. It should be reasonably high to avoid measurement issues in both waves, but it should not be too high, as there would not be enough remaining changes. I consider a threshold of exactly 10 minutes reasonable, that is: I identify the effect of commuting time on sick leave absences (or presenteeism) only for people with a change in commuting time equal to or higher than 10 minutes. The corresponding coefficients are "TiTrue" in Table 10 and in Table 11. For instance:  $TiTrue = Ti * 1(|Ti_{2016} - Ti_{2013}| \geq 10min)$ .<sup>30</sup> From this point on, the associated coefficients are my preferred estimates. Alternative thresholds of more than 5 minutes, more than 10 minutes or more than 15 minutes are also considered. Results are qualitatively unchanged. Neither does an attempt to test all possible thresholds between 0 minutes and 20 minutes (with 1-minute steps), as can be seen in Figure ?? and in Figure ?. No effect of commuting time on sick leave absences or on presenteeism is discernible.

In a robustness test, I also consider another specification for large changes, by which I mean a change larger than 5, 10 or 15 minutes.<sup>31</sup> I create a variable equal to 1 (respectively to -1) in 2016 if there is a large increase (respectively a large

---

<sup>30</sup>Alternatively, "TiSmall" designates the effect of commuting time on the dependent variable, as identified only on small changes. Both terms will be employed in all the subsequent tables with that meaning.  $TiSmall = Ti * 1(|Ti_{2016} - Ti_{2013}| < 10min)$

<sup>31</sup>Changes larger than 20 or 30 minutes were tested as well. Results are presented in Table H.II; they also yield non-significant results. It is however to be noted that only 10% of the sample (respectively 6% of the sample) experiences a change higher than 20 minutes between the two waves (resp. 30 minutes).

decrease) in commuting time between 2013 and 2016. The results are displayed in Table H.I of Annex H. All the results are also non-significant.

**TABLE 10.**  
Large changes in commuting time and sick leave absences

Specification	Absenteeism $A_i$				
	(1) Fixed-Effect Poisson all	(2) Fixed-Effect Poisson $ \Delta Ti  > 5\text{min}$	(3) Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(4) Fixed-Effect Poisson $ \Delta Ti  > 10\text{min}$	(5) Fixed-Effect Poisson $ \Delta Ti  > 15\text{min}$
obs. for TiTrue					
TiTrue	0.00112 (0.00182)	0.00108 (0.00185)	0.00130 (0.00189)	0.00114 (0.00185)	0.00115 (0.00196)
TiSmall		0.00247 (0.0110)	-0.00120 (0.00723)	0.000660 (0.0105)	0.000938 (0.00495)
Controls	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes
Observations	14206	14206	14206	14206	14206

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

**TABLE 11.**  
Large changes in commuting time and presenteeism

Specification	Presenteeism $P_i$				
	(1) Fixed-Effect Poisson	(2) Fixed-Effect Poisson	(3) Fixed-Effect Poisson	(4) Fixed-Effect Poisson	(5) Fixed-Effect Poisson
obs. for TiTrue	all	$ \Delta Ti  > 5\text{min}$	$ \Delta Ti  \geq 10\text{min}$	$ \Delta Ti  > 10\text{min}$	$ \Delta Ti  > 15\text{min}$
TiTrue	0.00198 (0.00219)	0.00153 (0.00221)	0.00195 (0.00225)	0.00168 (0.00221)	0.00161 (0.00231)
TiSmall		0.0191 (0.0148)	0.00250 (0.00896)	0.0123 (0.0143)	0.00486 (0.00677)
Controls	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes
Observations	17494	17494	17494	17494	17494

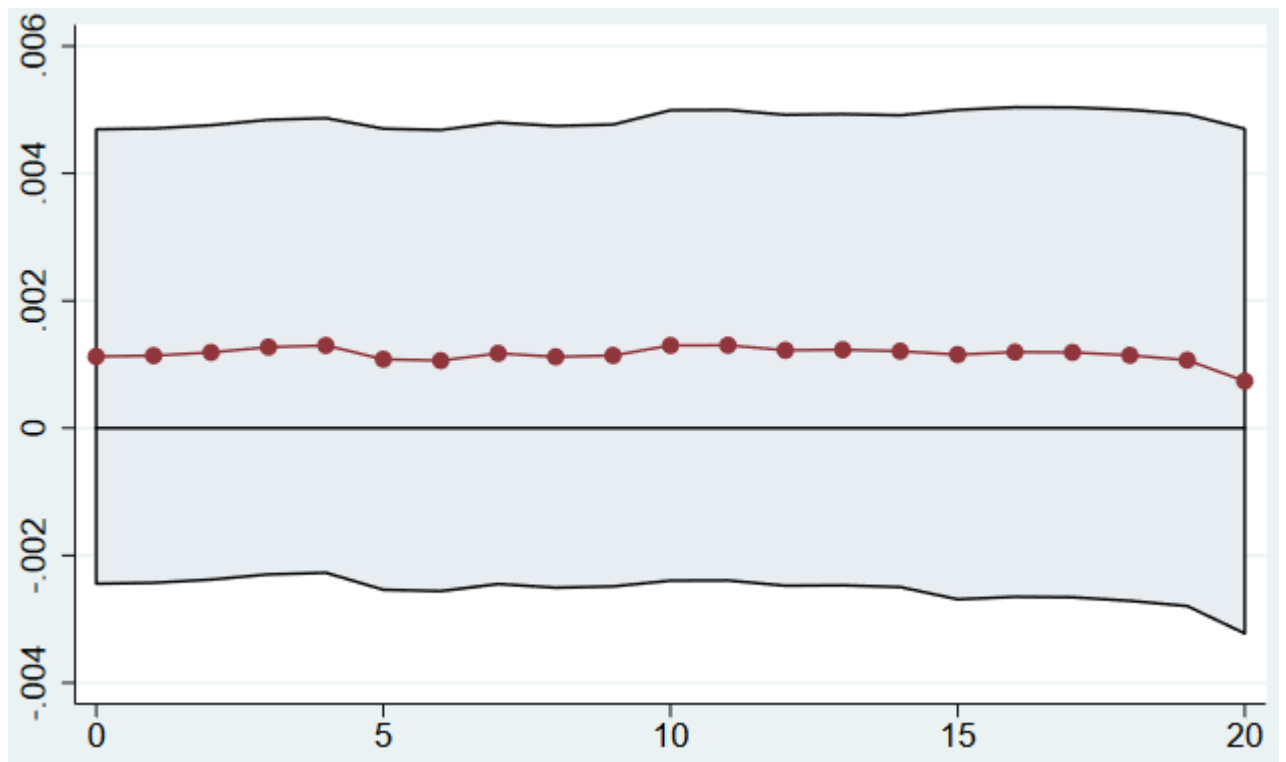
*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

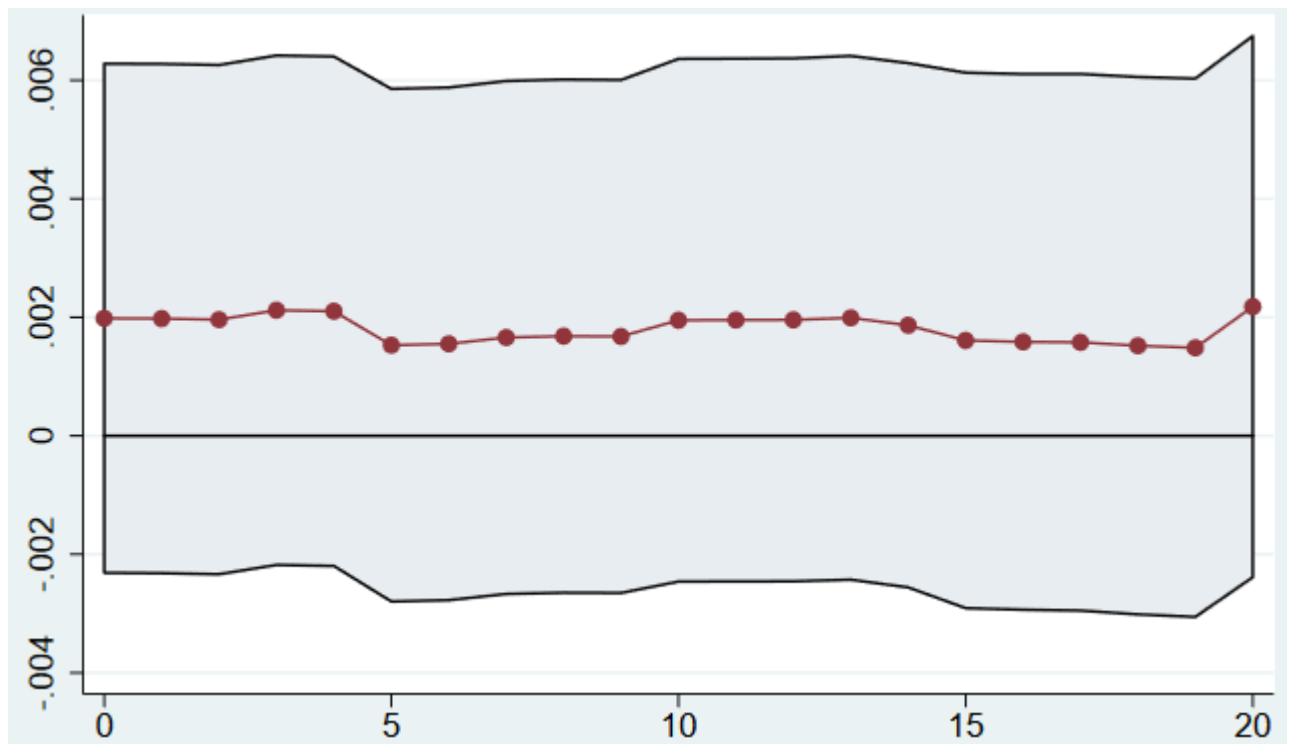
*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

**FIGURE 1.** Effect of Commuting Time  $T_i$  on sick leave absences, when estimated only based on changes in  $T_i$  between the two waves that are higher than a given threshold, for thresholds between 0 and 20 minutes.



Note: This is a graphical illustration of Table 10, with 1-minute precision. I present as well the 95% confidence interval.

**FIGURE 2.** Effect of Commuting Time  $T_i$  on Presenteeism, when estimated only based on changes in  $T_i$  between the two waves that are higher than a given threshold, for thresholds between 0 and 20 minutes.



Note: This is a graphical illustration of Table 11, with 1-minute precision. I present as well the 95% confidence interval.

## 4.5 Heterogeneity

I next test for heterogeneity in the impact of commuting time on absenteeism and presenteeism: first regarding gender and sectors (being in the public sector or being self-employed), and second regarding (log of) income, self-reported health, age, number of working hours and part-time work.

The results for gender and sectors are presented in Table 12 and in Table 13. They are my primary interest regarding heterogeneity for two reasons. First, the literature has sometimes found significant differences according to gender, and often investigated differences between sectors. Secondly, both require investigation due to the imbalance of my sample compared to the French employed population, mostly due to oversampling of the public sector.

In this section, I do not present in the tables the coefficients of the variables of interests, but the exponentiated coefficients (i.e. the exponential of the coefficient). The interpretation of the interactions in non-linear models requires special care as the coefficients in front of the interaction effects are not necessarily equal to the marginal effects (the cross derivatives) (Ai and Norton, 2003): this marginal effect depends on all the other explanatory variables, including the fixed effect, which I do not estimate to avoid the incidental parameter problem. That is why the various Stata commands commonly used in those circumstances (such as the `inteff` command of Norton, Wang and Ai (2004)) are not suited for models with fixed effects. Happily, an alternative approach, advocated by Buis (2010) can be applied in those cases. Acknowledging the problem raised by non-linear models, he notes that "*we can interpret interactions without referring to any additional program by presenting effects as multiplicative effects (for example, odds ratios, incidence-rate ratios, hazard ratios).*" The idea is to present the incidence rate ratios (i.e. the exponentiated coefficients), and correctly interpreting it as multiplicative effects.



Interpretation of the results presented in those tables remains simple. For instance, there is a 1.002 exponentiated coefficient for  $TiTrue$  in Column (2) of Table 12. It means that a 1-minute change in  $TiTrue$  multiplies the Expected number of days of absenteeism by 1.001. Furthermore, for men, due to the 0.997 exponentiated coefficient for  $TiTrue * male$ , the Expected number of days of absenteeism is also multiplied by 0.997. Both factors do not significantly differ from zero.

Considering these multiplicative effects, no interacted coefficient is significant for sick leave absences, which shows that there is no commuting-absences relationship specific to a gender or to the public sector. Regarding the self-employed, they are not part of my main sample, as their sick leave behavior might *a priori* differ from the rest of the active population. However, I do not find confirmation of this suspicion in my data, nor do I find a significant and distinct "self-employed effect" regarding their commuting-absence (absence of) relationship.

Similarly, no coefficient is significant when investigating possible heterogeneity of the commuting - presenteeism relationship, save for the interaction with gender. It is at the 10% confidence threshold, which generally implies that we cannot dismiss the possibility that there is such a relationship. Commuting time would appear to have a stronger effect on men than on women. I thus further investigate a possible effect related only to men in Table 14. Column (2) shows that the effect is also significant when estimated only on the men (at the 5% threshold). Column (3) shows this is robust considering all changes in commuting time (and not only the largest ones). But such a result is not robust to considering the Negative Binomial specification, as shown in Column (4). More importantly, it also disappears when I exclude the few observations with more than 3 months of presenteeism over the past 12 months.<sup>32</sup> As it is more likely that these few men have suffered a large health shock that has changed their commuting time, than

---

<sup>32</sup>When performing a grid search on this threshold in Figure 3, I find that the effect is no longer significant if I withdraw the 716 observations with more than 300 days of presenteeism.

the reverse,<sup>33</sup> this cannot be considered a robust result.

**TABLE 12.**  
Heterogeneity in the relation between commuting time and sick leave absences  
by gender and sector

Specification	Absenteeism $A_i$			
	(1) Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(2) Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(3) Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(4) Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$
TiTrue	1.001 (0.00185)	1.002 (0.00229)	1.001 (0.00236)	1.002 (0.00185)
TiTrue * male		0.997 (0.00385)		
TiTrue * public			1.001 (0.00324)	
TiTrue * independent				1.008 (0.0208)
independent				0.763 (0.294)
Controls	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes
Observations	14206	14206	14206	14574

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G). The Table presents the exponential of the variable coefficients. For instance, in model (1), the coefficient associated with TiTrue would be  $\ln(1.001)$ .

<sup>33</sup>Furthermore, as argued by Van Ommeren and Gutiérrez-i Puigarnau (2011) to justify dismissing their observations for which absenteeism was more than 30 days, estimates of count models are not consistent given random measurement error in the dependent variable. And such measurement errors might be particularly large for workers with a large number of days of presenteeism.

**TABLE 13.**  
Heterogeneity in the relation between commuting time and presenteeism by gender and sector

Specification	Presenteeism $P_i$			
	(1)	(2)	(3)	(4)
	Fixed-Effect Poisson obs. for TiTrue $ \Delta Ti  \geq 10\text{min}$	Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed-Effect Poisson $ \Delta Ti  \geq 10\text{min}$
TiTrue	1.002 (0.00221)	0.999 (0.00308)	1.003 (0.00260)	1.002 (0.00219)
TiTrue * male		1.007* (0.00413)		
TiTrue * public			0.996 (0.00400)	
TiTrue * independent				1.010 (0.00868)
independent				1.191 (0.269)
Controls	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes
Observations	17494	17494	17494	18926

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G). The Table presents the exponential of the variable coefficients. For instance, in model (1), the coefficient associated with TiTrue would be  $\ln(1.002)$ .

**TABLE 14.**  
Commuting-presenteeism relationship of men

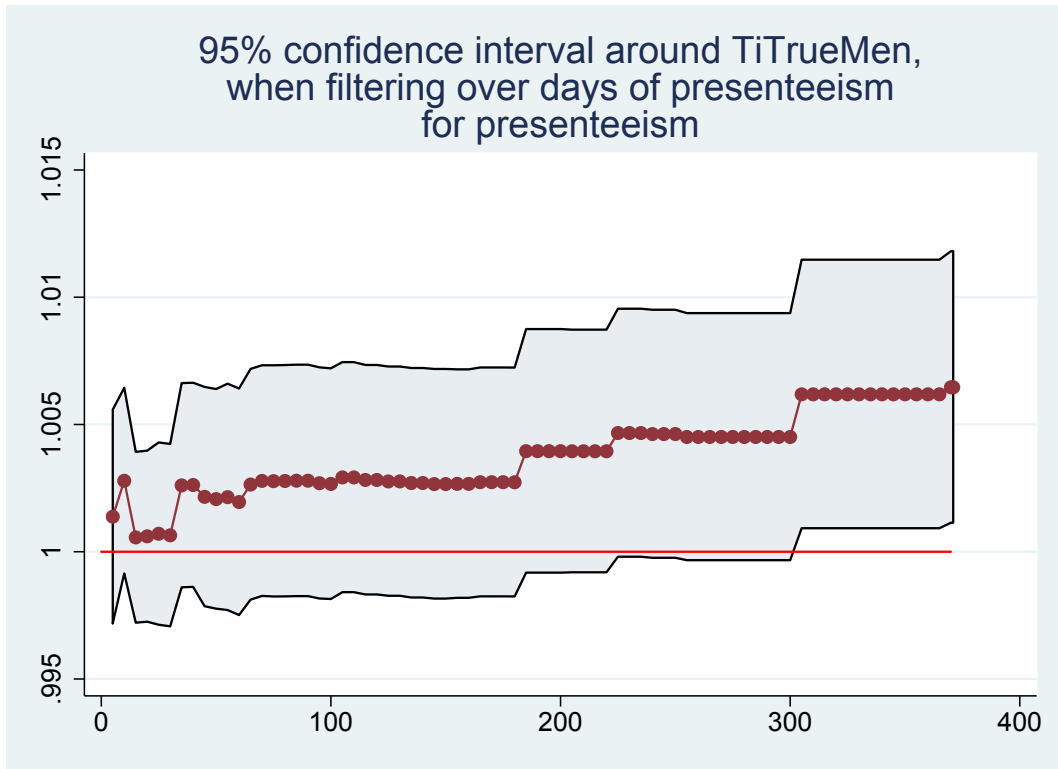
Specification	Presenteeism $P_i$				
	(1) Fixed Effect Poisson	(2) Fixed Effect Poisson	(3) Fixed Effect Poisson	(4) Fixed Effect Negative Binomial	(5) Fixed Effect Poisson
obs. for TiTrue restriction	$ \Delta Ti  \geq 10\text{min}$ -	$ \Delta Ti  \geq 10\text{min}$ -	all -	$ \Delta Ti  \geq 10\text{min}$ -	$ \Delta Ti  \geq 10\text{min}$ ( $P_i < 90$ )
TiTrue	0.999 (0.00308)				
TiTrue * male	1.007* (0.00413)	1.006** (0.00272)	1.007** (0.00276)	1.002 (0.00166)	1.003 (0.00232)
TiTrue * female		0.999 (0.00308)	1.000 (0.00304)	1.000 (0.00173)	1.000 (0.00211)
Controls	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes
Observations	17494	17494	17494	17494	17342

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered. For the Negative Binomial Regression (NegBin), the clustered robust standard errors are estimated through 50 bootstrap replications.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G). The Table presents the exponential of the variable coefficients. For instance, in model (1), the coefficient associated with TiTrue would be  $\ln(0.999)$ .



**FIGURE 3.** Grid search related to column (5) of Table 14

Legend: The x-axis shows the filtering condition. For instance, at  $x=90$ , all observations with a number of days of presenteeism  $P_i$  over 90 days are erased. The corresponding point is the coefficient of column (5) of Table 14. The point at the far right is the coefficient of column (2) of Table 14 (as no data can have more than 366 days of yearly presenteeism).

Estimates in which the focal variable is interacted with (log of) income, self-reported health, age, the number of working hours, the fact of working part time, of being a parent, or of having a child under 3 year old are presented in Tables 15 and 16. Such variables could be linked, for instance, to access to a doctor, as a medical certificate is necessary for all sick leave in France. However, no interaction is significant.

**TABLE 15.**  
Heterogeneity in the relation between commuting time and absenteeism

Specification	Absenteeism $A_i$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
obs. for TiTrue	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$	Fixed Effect Poisson $ \Delta Ti  \geq 10min$
TiTrue	1.001 (0.00202)	1.001 (0.00204)	1.001 (0.00202)	1.001 (0.00202)	1.001 (0.00202)	1.001 (0.00202)	1.001 (0.00202)
TiTrue * wage	0.998 (0.00407)						
TiTrue * good health		1.001 (0.00228)					
TiTrue * age			1.000 (0.000578)				
TiTrue * working hours				1.000 (0.000265)			
TiTrue * part time					0.994 (0.00586)		
TiTrue * children						1.000 (0.00541)	
TiTrue * children_under_3							0.995 (0.00486)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12392	12392	12392	12392	12392	12392	12392

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Controls include personal-related variables (a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies). "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

**TABLE 16.**  
Heterogeneity in the relation between commuting time and presenteeism

Specification	Presenteeism $P_i$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
obs. for TiTrue	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$
TiTrue	1.002 (0.00249)	1.002 (0.00236)	1.002 (0.00249)	1.002 (0.00249)	1.002 (0.00248)	1.002 (0.00247)	1.002 (0.00249)
TiTrue * wage	1.004 (0.00301)						
TiTrue * good health		1.003 (0.00247)					
TiTrue * age			0.999 (0.000732)				
TiTrue * working hours				1.000 (0.000279)			
TiTrue * part time					1.000 (0.00539)		
TiTrue * children						1.000 (0.00879)	
TiTrue * children_under_3							0.994 (0.00513)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15334	15334	15334	15334	15334	15334	15334

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Controls include personal-related variables (a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies). "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).



## **4.6 Dealing with weights**

Overall, the public sector is notably over-represented in my sample. As such, I identify an average effect within a population which is not representative of my true population of interest (the French active population). Even though I do not find any sector-related effect in the commuting-sick leave absences (or commuting - presenteeism) relationship, there could be hidden heterogeneity that may make my estimates differ from the true average effect in the French active population (Solon, Haider and Wooldridge, 2015). Panel weights are provided to make the survey representative of the French active population in 2013. When applying them in the fixed-effect linear regression (for which `pweights/aweight`s options are available, unlike in fixed-effects Poisson or Negative Binomial regressions), results remain non-significant.

**TABLE 17.**  
With and without weights

Specification	Absenteeism $A_i$			Presenteeism $P_i$		
	(1) Fixed Effect Poisson $ \Delta T_i  \geq 10\text{min}$ no	(2) Fixed Effect linear $ \Delta T_i  \geq 10\text{min}$ no	(3) Fixed Effect linear $ \Delta T_i  \geq 10\text{min}$ yes	(4) Fixed Effect Poisson $ \Delta T_i  \geq 10\text{min}$ no	(5) Fixed Effect linear $ \Delta T_i  \geq 10\text{min}$ no	(6) Fixed Effect linear $ \Delta T_i  \geq 10\text{min}$ yes
obs. for TiTrue weights						
TiTrue	0.00114 (0.00185)	0.00688 (0.00733)	0.00599 (0.0119)	0.00168 (0.00221)	0.00399 (0.00831)	0.00962 (0.00862)
TiSmall	0.000660 (0.0105)	0.00848 (0.0448)	0.0845 (0.0705)	0.0123 (0.0143)	0.0314 (0.0439)	-0.0112 (0.0460)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14206	30552	30552	17494	29696	29696

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available at the two periods. The Poisson specification also does not take into account all observations with a null value for the two periods.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

## **4.7 Testing another focal variable**

I have previously investigated the possible impact of more commuting time on sick leave (or presenteeism). An implicit hypothesis behind the fixed-effect specification is that a 10-minute change between two waves has the same impact if it increases commuting time to 10 minutes or to 1 hour. Although this is what has been commonly tested in the literature, I am aware that Goerke and Lorenz (2017) showed that only long commuters endure more sick leave due to their commute. What matters may not be "spending 20 more minutes commuting", but "spending (for instance) more than 45 minutes commuting".

I test the following hypothesis in Tables 18 and 19. First, I consider three dummies: commuting time between 5 minutes and 24 minutes; between 25 minutes and 44 minutes; and above 45 minutes. People in the reference category are thus those with a negligible commute (under 5 minutes). The 24-minute threshold has been chosen as it is the mean commuting time in my sample (see Table 2). Secondly, I consider only one dummy: being above 45 minutes. Here, People in the reference category are those with a commuting time under 44 minutes. Thirdly, I consider only one dummy: being above 24 minutes. Here People in the reference category are those with a commuting time under the mean commuting time.

For both absenteeism and presenteeism, very long commutes statistically increase the dependent variable in pooled regressions. All significance vanished, however, as soon as the fixed effect are introduced into the regression, that is as soon as I control for the individual time-unvarying unobserved characteristics. Unlike Goerke and Lorenz (2017) I therefore do not find evidence to support the idea that the magnitude of the commute matters.

**TABLE 18.**  
Alternative focal variable that may cause absenteeism

Specification	Absenteeism $A_i$					
	(1) Poisson	(2) Poisson	(3) Poisson	(4) Fixed Effect Poisson	(5) Fixed Effect Poisson	(6) Fixed Effect Poisson
Ti_d5_24	0.0476 (0.0741)			-0.0239 (0.126)		
Ti_d25_44	0.0944 (0.0781)			0.0399 (0.135)		
Ti_d45_more	0.143* (0.0840)	0.0830* (0.0459)		-0.0711 (0.143)	-0.0906 (0.0893)	
Ti_d25_more			0.0663** (0.0324)			0.0379 (0.0703)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30777	30777	30777	14206	14206	14206
fixed-effects	No	No	No	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).  $Ti\_d5\_24$  is a dummy for having a commuting time Ti between 5 and 24 minutes (included).

**TABLE 19.**  
Alternative focal variable that may cause presenteeism

Specification	Presenteeism $P_i$					
	(1) Poisson	(2) Poisson	(3) Poisson	(4) Fixed Effect Poisson	(5) Fixed Effect Poisson	(6) Fixed Effect Poisson
Ti_d5_24	0.0476 (0.0741)			-0.0239 (0.126)		
Ti_d25_44	0.0944 (0.0781)			0.0399 (0.135)		
Ti_d45_more	0.143* (0.0840)	0.0830* (0.0459)		-0.0711 (0.143)	-0.0906 (0.0893)	
Ti_d25_more			0.0663** (0.0324)			0.0379 (0.0703)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30777	30777	30777	14206	14206	14206
fixed-effects	No	No	No	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).  $Ti\_d5\_24$  is a dummy for having a commuting time Ti between 5 and 24 minutes (included).

## 5 The intermediation of health in the studied relationships

### 5.1 Psychological well-being may intermediate most of the studied relationships

I now replicate Table 6 and Table 7 with my improved specification: fixed-effect regressions with *TiTrue* as a focal variable and with all the controls I added to take the possible endogenous choices of new workplace or place of residence into account (see Table 20 and Table 21). Except for one coefficient that becomes significant at the 10% threshold, all coefficients remain non-significant. The main purpose of these regressions is however to investigate whether adding self-assessed health and psychological well-being to the regressions decrease the coefficients of interest, as previously observed for presenteeism and sometimes observed for absenteeism.

The corresponding changes in the coefficients can be computed by comparing the *TiTrue* coefficient between Columns (3) and Columns (2) (and between Columns (6) and Columns (5) for the Negative Binomial specification) of Table 6 (and of Table 7 for presenteeism). The corresponding differences are also computed in Table C.IV, where their statistical significance is furthermore tested. I do observe a decrease in coefficients of interest whenever I add self-assessed health and psychological well-being to the regressions. However, these decreases are relatively small and non-significant in the case of absenteeism, but strong and significant in the case of presenteeism. If the relationships under study actually exist and I am just lacking statistical power to uncover them, this suggests that part of such relationships might be intermediated by self-assessed health or psychological well-being, particularly for presenteeism.

This hypothesis is further supported by the fact that these intermediating variables - self-assessed health and psychological well-being - influences absenteeism

and presenteeism strongly and significantly. For instance, a 1-point increase on the 5-point scale of self-assessed health drives a decrease in absenteeism and in presenteeism of approximately 35% (see columns (3) of Table 6 and Table 7). And a 1-point increase on the 25-point scale of psychological well-being drives decreases of roughly 2% and 4% in absenteeism and in presenteeism - respectively. This can be rephrased in terms of standard deviation knowing that one standard deviation in self-assessed health is 0.8 and one standard deviation in psychological well-being is 5. Hence, one standard deviation in self-assessed health changes absenteeism and presenteeism by approximately 28%, while one standard deviation in psychological well-being changes absenteeism by approximately 10% and presenteeism by approximately 20%.

In order to substantiate the hypothesis that commuting time may impact absenteeism and/or presenteeism through changes in health, one would first need to show that an increase in commuting time leads to a decrease in self-assessed health and/or psychological well-being. This is investigated in Section 5.2.

**TABLE 20.**  
Use of fixed-effect models to study the impact of commuting time on absenteeism

Specification	Absenteeism $A_i$					
	(1) Fixed Effect Poisson	(2) Fixed Effect Poisson	(3) Fixed Effect Poisson	(4) Fixed Effect Negative Binomial	(5) Fixed Effect Negative Binomial	(6) Fixed Effect Negative Binomial
TiTrue	0.000682 (0.00202)	0.000888 (0.00203)	0.000867 (0.00202)	0.00270 (0.00345)	0.00269 (0.00281)	0.00197* (0.00110)
TiSmall	0.000129 (0.0112)	0.000811 (0.0112)	0.00374 (0.0111)	-0.000723 (0.00312)	-0.000822 (0.00457)	-0.00134 (0.00195)
log(monthly income)		-0.167 (0.102)	-0.132 (0.103)		0.00933 (0.360)	0.0470 (0.0654)
working hours		0.0121** (0.00554)	0.0121** (0.00543)		0.00599 (0.00954)	0.00454 (0.00279)
good health			-0.328*** (0.0359)			-0.305*** (0.0266)
psychological well-being			-0.0223*** (0.00591)			-0.0233*** (0.00426)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12392	12392	12392	12392	12392	12392
Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered. For the Negative Binomial Regression (NegBin), the clustered robust standard errors are estimated through 50 bootstrap replications.



**TABLE 21.**  
Use of fixed-effect models to study the impact of commuting time on  
presenteeism

Specification	Absenteeism $A_i$					
	(1) Fixed Effect Poisson	(2) Fixed Effect Poisson	(3) Fixed Effect Poisson	(4) Fixed Effect Negative Binomial	(5) Fixed Effect Negative Binomial	(6) Fixed Effect Negative Binomial
TiTrue	0.00181 (0.00248)	0.00210 (0.00250)	0.000654 (0.00220)	0.000557 (0.000812)	0.000538 (0.000841)	-0.000157 (0.00184)
TiSmall	0.0131 (0.0153)	0.0142 (0.0153)	0.0193 (0.0157)	0.00282 (0.00177)	0.00280 (0.00179)	0.00200 (0.00189)
log(monthly income)		0.0600 (0.0908)	0.0731 (0.0969)		-0.0133 (0.0480)	0.0324 (0.0481)
working hours		0.0172** (0.00681)	0.0137** (0.00641)		0.0121*** (0.00305)	0.0106*** (0.00368)
good health			-0.366*** (0.0454)			-0.199*** (0.0288)
psychological well-being			-0.0450*** (0.00944)			-0.0458*** (0.00361)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15334	15334	15334	15334	15334	15334
Fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered. For the Negative Binomial Regression (NegBin), the clustered robust standard errors are estimated through 50 bootstrap replications.

## 5.2 Does commuting time have a direct health effect?

Before investigating the effect of commuting time on self-assessed health and psychological well-being, it should be noted that the literature has found that commuting decreases the psychological health of women (but not men) in the UK (Roberts, Hodgson and Dolan (2011), Dickerson, Hole and Munford (2014)).

I investigate the effect of commuting time on these variables by using alternatively a fixed-effect linear regression, a fixed-effect Poisson regression (as it is a count variable in a sense) and a fixed-effect ordered logit<sup>34</sup> (as it comes from ordered variables). Note that in their methodological paper on the relationship between well-being and commuting, Dickerson, Hole and Munford (2014) used both a linear fixed-effect model and (various) fixed-effect ordered logit models to model a similar Likert scale variable. Once more, my preferred specification for non-negative integers remains the fixed-effect Poisson model for its robustness properties (see Annex B).

Table 22 presents the results on self-assessed health. All the coefficients of interests are positive but most of them are not significant. They are significant for three specifications out of three when specifically looking at an effect on women. In those cases, they are significant at the 10% threshold, which can only be interpreted as a sign that there might be an effect, rather than evidence that there is an effect. Moreover, the corresponding magnitude is rather low: considering column (5), a 20-minute increase in commuting time would imply a decrease by 0.8% in the self-assessed health score for women. With an average female self-assessed health of 2.12, this would imply a decrease by 0.2 point on the 5-point scale.

Besides, in all the models in Table 23, I find a negative and significant effect of commuting time on psychological well-being. As in the UK (Roberts, Hodgson

---

<sup>34</sup>I use the user-written Stata command `feologit_buc` implemented by Baetschmann, Staub and Winkelmann (2015). It was retrieved on: <https://rss.onlinelibrary.wiley.com/doi/10.1111/rssa.12166>.

and Dolan (2011)), Dickerson, Hole and Munford (2014)) and similarly to Table 22, this is driven by women. The corresponding magnitude also seems quite low: a 20-minute increase in commuting time implies a decrease of 0.8% in the psychological well-being for the sample (see column 2); with a mean psychological well-being of 16 points, this implies an increase of only 0.13 points on the 25-point scale. When focusing on women, the corresponding increases are of 1.2%, that is on average 0.19 points on the 25-point scale. Despite the large impact of psychological well-being on absenteeism and presenteeism, these (significant) effects of commuting time on psychological well-being are quite small, and are thus unlikely to cause subsequently large increases in neither absenteeism and presenteeism.

**TABLE 22.**  
Effect of commuting time on self-assessed health

Specification	Mean effect on health			Effect on health by gender		
	(1) Fixed Effect linear $ \Delta Ti  \geq 10\text{min}$	(2) Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(3) Fixed Effect ordered logit $ \Delta Ti  \geq 10\text{min}$	(4) Fixed Effect linear $ \Delta Ti  \geq 10\text{min}$	(5) Fixed Effect Poisson $ \Delta Ti  \geq 10\text{min}$	(6) Fixed Effect ordered logit $ \Delta Ti  \geq 10\text{min}$
obs. for TiTrue						
TiTrue	-0.000609 (0.000440)	-0.000210 (0.000150)	-0.00248 (0.00167)			
TiSmall	0.00307 (0.00243)	0.00107 (0.000831)	0.00791 (0.00956)	0.00307 (0.00243)	0.00107 (0.000831)	0.00792 (0.00956)
TiTrue * Man				-0.000108 (0.000601)	-0.0000412 (0.000198)	-0.000671 (0.00234)
TiTrue * Woman				-0.00111* (0.000639)	-0.000391* (0.000224)	-0.00419* (0.00235)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30942	30934	15740	30942	30934	15740

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the dependent variable, the focal variable (commuting time) and controls are available at the two periods. The Poisson specification also does not take into account all observations with a null dependent variable for the two periods.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

**TABLE 23.**  
Effect of commuting time on psychological well-being

Specification	Mean effect on well-being			Effect on well-being by gender		
	(1) Fixed Effect linear	(2) Fixed Effect Poisson	(3) Fixed Effect ordered logit	(4) Fixed Effect linear	(5) Fixed Effect Poisson	(6) Fixed Effect ordered logit
obs. for TiTrue	$ \Delta Ti  \geq 10min$	$ \Delta Ti  \geq 10min$	$ \Delta Ti  \geq 10min$	$ \Delta Ti  \geq 10min$	$ \Delta Ti  \geq 10min$	$ \Delta Ti  \geq 10min$
TiTrue	-0.00606** (0.00287)	-0.000396** (0.000185)	-0.00340** (0.00151)			
TiSmall	0.0167 (0.0157)	0.00108 (0.00100)	0.00694 (0.00873)	0.0167 (0.0157)	0.00108 (0.00100)	0.00696 (0.00873)
TiTrue * Man				-0.00248 (0.00379)	-0.000161 (0.000238)	-0.00164 (0.00217)
TiTrue * Woman				-0.00947** (0.00426)	-0.000632** (0.000282)	-0.00483** (0.00211)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28040	28038	101194	28040	28038	101194

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the dependent variable, the focal variable (commuting time) and controls are available at the two periods. The Poisson specification also does not take into account all observations with a null dependent variable for the two periods.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm size-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. Standard-errors are robust and clustered.

*Legend:* "ChangeVariables" stands for dummies controlling for a change in place of residence or job in 2016. "AdditChangeVariables" stands for additional variables of the same kind (see Annex G).

## 6 Conclusion

Although the pooled regressions would imply that commuting time increases both sick leave and presenteeism, this correlation disappears when taking time-invariant characteristics into account. To my knowledge, it is the first time that a possible commuting-presenteeism relationship has been investigated. At the same time, the non-significant finding regarding sick leave absenteeism is at odds with previous studies on the impact of commuting on sick leave absences in Germany, the UK and the US. That this relationship might be country-dependent is reasonable, as both Goerke and Lorenz (2017) and Van Ommeren and Gutiérrez-i Puigarnau (2011) found an effect on both genders in Germany, but Gimenez-Nadal, Molina and Velilla (2018) found commuting increases absenteeism only for male US workers and Künn-Nelen (2016) found commuting decreases absenteeism only for British women.

Not finding any significant impact does not necessarily imply that there is no underlying relationship. Although I can discard that there are large effects, I may simply be lacking statistical power to detect smaller effects. For instance, in line with previous evidence in England, I am able to establish a significant impact of commuting time on psychological well-being, particularly for women. As I also find that a decrease in psychological well-being significantly increases absenteeism and presenteeism, commuting time should be expected to have an impact on absenteeism and presenteeism, at least through psychological well-being. Future examinations of the same data, completed with the coming waves of the French Working Condition Survey, may prove more conclusive thanks to greater statistical power.

However both relationships, if they exist, seem to be somewhat small in magnitude, especially for sick leave. My preferred estimates point to a 1-minute elasticity of absenteeism of 0.001, and to a 1-minute elasticity of presenteeism of 0.002. That is, one standard deviation in commuting time (20 minutes, which is also close to

the mean commuting time) would translate only into 2% more sick leave and 4% more presenteeism. This assertion is somewhat reinforced by my examination of two health variables, with self-assessed health and psychological well-being. For both variables, I find that commuting time would seem to have a small detrimental health effect, although they are strongly significant only for psychological well-being. Hence, if health effects are the main causal channel between commuting time and absenteeism/presenteeism, large effects should not be expected on these latter variables as psychological well-being and self-assessed health are not hugely affected.

Though it is not the main object of this paper, it might be worth mentioning that such impacts of commuting time on health, especially on psychological well-being, are interesting results *per se*. First they hint that the well-being of (female) commuters might be directly affected, which is a further evidence of the "commuter paradox" first highlighted by (Stutzer and Frey, 2008).<sup>35</sup> Secondly, that these effects are born by the women coincide with the other articles that focus on those variables of interest that I am aware on (on British data; see Roberts, Hodgson and Dolan (2011), Dickerson, Hole and Munford (2014), Munford et al. (2018)). These gender differences could in turn limit the professional mobility of women, and contribute to the gender gap. I would think there is an avenue for research investigating the reasons behind such differences between genders.

Finally, this paper provides a few words of caution for other existing research papers studying any impact of commuting. The strategy consisting in keeping only those people who declare that they have changed neither place of residence nor job may be misleading. Unlike what has been stated by some authors, it does not seem to make the identification come primarily from people whose firm moves. In my data, few people seem concerned by such an exogenous change. Rather, among

---

<sup>35</sup>The commuter paradox is that "According to standard economics, the burden of commuting is chosen when compensated either on the labor or on the housing market so that individuals' utility is equalized." (Stutzer and Frey, 2008) But this is contradicted when people with longer commuting time report lower subjective well-being.

other plausible explanations, some people have jobs that imply regular changes in the commuting time (distance) ; or true changes in job or place of residence are hard to identify in the data. This does not bring into question the need to take possible time-varying changes that may be endogenous into account as much as possible. But this shows that having an even cleaner identification strategy based on firms' relocations requires actual information on those relocations (and, apart from event studies, a very large database due to their rarity). To my knowledge, in the literature on the impacts of commuting, this has only been done by Mulalic, Van Ommeren and Pilegaard (2013) to study the impact of commuting on wages using (exhaustive) Danish administrative data.



## References

- Ai, Chunrong, and Edward C Norton.** 2003. “Interaction terms in logit and probit models.” *Economics letters*, 80(1): 123–129.
- Allison, Paul D, and Richard P Waterman.** 2002. “Fixed-effects negative binomial regression models.” *Sociological methodology*, 32(1): 247–265.
- Angrist, Joshua D, and Jörn-Steffen Pischke.** 2008. *Mostly harmless econometrics: An empiricist’s companion*. Princeton university press.
- Arnold, Daniel, Marco de Pinto, et al.** 2015. “How are Work-related Characteristics Linked to Sickness Absence and Presenteeism? Theory and Data.” *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, 135(4): 465–498.
- Baetschmann, Gregori, Kevin E Staub, and Rainer Winkelmann.** 2015. “Consistent estimation of the fixed effects ordered logit model.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3): 685–703.
- Beland, Louis-Philippe, and Daniel A Brent.** 2018. “Traffic and crime.” *Journal of Public Economics*, 160: 96–116.
- Bergström, Gunnar, Lennart Bodin, Jan Hagberg, Gunnar Aronsson, and Malin Josephson.** 2009. “Sickness presenteeism today, sickness absenteeism tomorrow? A prospective study on sickness presenteeism and future sickness absenteeism.” *Journal of Occupational and Environmental Medicine*, 51(6): 629–638.
- Bolusset, Armelle, and Christophe Rafrac.** 2019. “Sept salariés sur dix vont travailler en voiture.” Institut National de la Statistique et des Etudes Economiques.
- Buis, Maarten L.** 2010. “Stata tip 87: Interpretation of interactions in nonlinear models.” *The stata journal*, 10(2): 305–308.

- Cazenave-Lacroutz, Alexandre, and Alexandre Godzinski.** 2017. “Effects of the one-day waiting period for sick leave on health-related absences in the French central civil service.” Institut National de la Statistique et des Etudes Economiques.
- Chatterji, Monojit, and Colin J Tilley.** 2002. “Sickness, absenteeism, presenteeism, and sick pay.” *Oxford Economic Papers*, 54(4): 669–687.
- Clogg, Clifford C, Eva Petkova, and Adamantios Haritou.** 1995. “Statistical methods for comparing regression coefficients between models.” *American Journal of Sociology*, 100(5): 1261–1293.
- Coudène, Maud, and David Levy.** 2016. “De plus en plus de personnes travaillent en dehors de leur commune de résidence.” Institut National de la Statistique et des Etudes Economiques.
- Delance, Pauline, and Benjamin Vignolles.** 2017. “Ça déménage ? La mobilité résidentielle et ses déterminants.” Institut National de la Statistique et des Etudes Economiques.
- Dickerson, Andy, Arne Risa Hole, and Luke A Munford.** 2014. “The relationship between well-being and commuting revisited: Does the choice of methodology matter?” *Regional Science and Urban Economics*, 49: 321–329.
- Gallais, Lenka, and Michael J Griffin.** 2006. “Low back pain in car drivers: a review of studies published 1975 to 2005.” *Journal of sound and vibration*, 298(3): 499–513.
- Gimenez-Nadal, J Ignacio, José Alberto Molina, and Jorge Velilla.** 2018. “Commuting Time and Sick-Day Absence of US Workers.” *IZA Discussion Papers*.
- Goerke, Laszlo, and Olga Lorenz.** 2017. “Commuting and sickness absence.” *mimeo*.

- Goetzl, Ron Z, Stacey R Long, Ronald J Ozminkowski, Kevin Hawkins, Shaohung Wang, and Wendy Lynch.** 2004. "Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers." *Journal of Occupational and Environmental Medicine*, 46(4): 398–412.
- Gould, William.** 2011. "Use poisson rather than regress; tell a friend." *The Stata blog*, accessed online on the 11st of July 2019.
- Gutiérrez-i Puigarnau, Eva, and Jos N van Ommeren.** 2010. "Labour supply and commuting." *Journal of Urban Economics*, 68(1): 82–89.
- Gutiérrez-i Puigarnau, Eva, and Jos N van Ommeren.** 2015. "Commuting and labour supply revisited." *Urban Studies*, 52(14): 2551–2563.
- Hausman, Jerry, Bronwyn H Hall, Zvi Griliches, et al.** 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica*, 52(4): 909–938.
- Inan, Ceren.** 2020. "Quel lien entre les conditions de travail et le présentéisme des salariés en cas de maladie?"
- Johansson, Per, and Mårten Palme.** 2005. "Moral hazard and sickness insurance." *Journal of Public Economics*, 89(9): 1879–1890.
- Kaiser, Boris.** 2015. "RHAUSMAN: Stata module to perform robust Hausman specification test."
- Künn-Nelen, Annemarie.** 2016. "Does commuting affect health?" *Health economics*, 25(8): 984–1004.
- Lancaster, Tony.** 2000. "The incidental parameter problem since 1948." *Journal of econometrics*, 95(2): 391–413.
- Lorenz, Olga.** 2018. "Does commuting matter to subjective well-being?" *Journal of Transport Geography*, 66: 180–199.

- Mulalic, Ismir, Jos N Van Ommeren, and Ninette Pilegaard.** 2013. "Wages and commuting: Quasi-natural experiments' evidence from firms that relocate." *The Economic Journal*, 124(579): 1086–1105.
- Munford, Luke, Nigel Rice, Jennifer Roberts, and Nikita Jacob.** 2018. "The disutility of commuting? The effect of gender and local labour markets." *The Sheffield Economic Research Paper Series (SERPS)*, 201801(2018010).
- Neyman, Jerzy, Elizabeth L Scott, et al.** 1948. "Consistent estimates based on partially consistent observations." *Econometrica*, 16(1): 1–32.
- Norton, Edward C, Hua Wang, and Chunrong Ai.** 2004. "Computing interaction effects and standard errors in logit and probit models." *The Stata Journal*, 4(2): 154–167.
- Pichler, Stefan, and Nicolas R Ziebarth.** 2017. "The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior." *Journal of Public Economics*, 156: 14–33.
- Pouliakas, Konstantinos, and Ioannis Theodossiou.** 2013. "The economics of health and safety at work: an interdisciplinary review of the theory and policy." *Journal of Economic Surveys*, 27(1): 167–208.
- Ree, Eline, Stein Atle Lie, Hege R Eriksen, Kirsti Malterud, Aage Indahl, Oddrun Samdal, and Anette Harris.** 2016. "Reduction in sick leave by a workplace educational low back pain intervention: a cluster randomized controlled trial." *Scandinavian journal of public health*, 44(6): 571–579.
- Roberts, Jennifer, Robert Hodgson, and Paul Dolan.** 2011. "'It's driving her mad": Gender differences in the effects of commuting on psychological health." *Journal of health economics*, 30(5): 1064–1076.
- Ross, Stephen L, and Yves Zenou.** 2008. "Are shirking and leisure substitutable? An empirical test of efficiency wages based on urban economic theory." *Regional Science and Urban Economics*, 38(5): 498–517.

- Solon, Gary, Steven J Haider, and Jeffrey M Wooldridge.** 2015. “What are we weighting for?” *Journal of Human resources*, 50(2): 301–316.
- Stutzer, Alois, and Bruno S Frey.** 2008. “Stress that doesn’t pay: The commuting paradox.” *Scandinavian Journal of Economics*, 110(2): 339–366.
- Topp, Christian Winther, Søren Dinesen Østergaard, Susan Søndergaard, and Per Bech.** 2015. “The WHO-5 Well-Being Index: a systematic review of the literature.” *Psychotherapy and psychosomatics*, 84(3): 167–176.
- Van Ommeren, Jos N, and Eva Gutiérrez-i Puigarnau.** 2011. “Are workers with a long commute less productive? An empirical analysis of absenteeism.” *Regional Science and Urban Economics*, 41(1): 1–8.
- Wooldridge, Jeffrey M.** 1999a. “Quasi-likelihood methods for count data.” *Handbook of applied econometrics*, 2.
- Wooldridge, Jeffrey M.** 1999b. “Distribution-free estimation of some nonlinear panel data models.” *Journal of Econometrics*, 90(1): 77–97.
- Yan, Jun, Robert H Aseltine Jr, and Ofer Harel.** 2013. “Comparing regression coefficients between nested linear models for clustered data with generalized estimating equations.” *Journal of Educational and Behavioral Statistics*, 38(2): 172–189.

# Appendices

## A Variables and sample composition

TABLE A.I. Variable definitions.

<i>Variable</i>	<i>Definition</i>
<b>Main variables</b>	
Sick leave absences $A_i$	Reported number of days on sick leave in the last twelve months. 0 if no sick leave episode.
Presenteeism $P_i$	Reported number of days of presenteeism in the last twelve months. 0 if no presenteeism episode.
Commuting time $T_i$	Time in minutes to make the outward trip from home to work, on average.
<b>core controls</b>	
<i>Personal characteristics</i>	
Male	==1 if the respondent is a male.
Age	Age in years.
Age <sup>2</sup>	Age in square.
In_couple	== 1 if the respondent has a partner.
Children	Children == 1 if there is any child in the household.
Children_under_3	Children_under_3 == 1 if there is any child under 3 year old in the household.
Educational dummies	5 dummies respectively equal to 1: if no diploma or at most the primary school certificate ; if at most a secondary school diploma ; if at most the certificate of professional competence (skilled manual worker or employee) ; if at most the baccalaureate ; in case of any higher education diploma.
<i>Job related characteristics</i>	
Tenure	Numbers of months with the same employer. The variable is computed by the survey producer based on the start of the contract. Due to the imprecision of this latest variable, I compute it for those not changing job between 2013 and 2016, based on the 2013 tenure to which I add the age difference between the two waves (times 12).
Tenure <sup>2</sup>	Tenure in square.
Part_time	Part_time==1 if part time work.
Permanent	Permanent==1 if in permanent contract or a civil servant.
Public	Public==1 if in the public sector.
Hospital	Hospital==1 if in the hospital (or long-term care) sector.
Public_hospital	Public_hospital==1 if in the public hospital (or long-term care) sector.
Independent	Independent==1 if self-employed (when self-employed are included).
Blue_collar	Blue_collar==1 if technician, skilled or unskilled worker.
Firm size dummies	9 dummies respectively equal to 1 if the number of employees in the firm is: between 1 and 4; between 5 and 9; between 10 and 19; between 20 and 49; between 50 and 199; between 200 and 499; between 500 and 999; 1000 or more.
Firm activity dummies	17 dummies corresponding from the 2008 French Statistical Classification of Economic Activities (NAF17v2), with 17 levels.
Area dummies	13 dummies corresponding to the 13 French regions.
<b>Additional co-variables</b>	
log(monthly income)	Current monthly income, net from social contributions. When it is provided within a bracket, I consider the mean of the bracket.
Working hours	usual number of working hours during a week.
good health (bsante)	An indicator of self-reported health status: 1 = "very poor"; 2 = "poor"; 3 = "acceptable"; 4 = "good"; 5 = "very good". ( <i>I reversed the original scale of bsante.</i> )
psychological well-being (who)	A 25-point Likert scale depicting subjective psychological well-being. It is based on questions regarding how often 5 feelings occur: "I felt good and in a good mood"; "I felt calm and quiet"; "I felt full of energy and vigorous"; "I woke up feeling fresh and refreshed"; "My daily life has been filled with interesting things". Answer to each questions could be: 1. All the time, 2. Most of the time, 3. More than half of the time, 4. Less than half of the time, 5. Occasionally, 6. Never, 9. Don't Know. It is reversed so that a higher score implies a higher psychological well-being.

**TABLE A.II.**  
Sample composition

	All		Men		Women	
	mean	sd	mean	sd	mean	sd
male	0.43	0.49	1	0	0	0
age	43.9	9.66	43.7	9.73	44.1	9.61
In_couple	0.77	0.42	0.80	0.40	0.75	0.43
children	0.63	0.48	0.61	0.49	0.64	0.48
children_under_3	0.11	0.31	0.12	0.32	0.095	0.29
at most the primary school certificate	0.080	0.27	0.094	0.29	0.070	0.26
at most a secondary school diploma	0.045	0.21	0.046	0.21	0.045	0.21
at most the certificate of professional competence	0.25	0.43	0.30	0.46	0.21	0.41
at most the baccalaureate	0.18	0.39	0.17	0.38	0.19	0.39
any higher education diploma	0.44	0.50	0.39	0.49	0.48	0.50
tenure	219.3	135.3	213.6	134.8	223.5	135.6
part_time	0.18	0.39	0.048	0.21	0.29	0.45
public	0.45	0.50	0.36	0.48	0.51	0.50
hospital	0.18	0.38	0.082	0.27	0.25	0.43
public_hospital	0.12	0.33	0.060	0.24	0.17	0.38
blue_collar	0.21	0.41	0.34	0.47	0.12	0.32
< 5 employees	0.061	0.24	0.057	0.23	0.063	0.24
5 - 9 employees	0.080	0.27	0.079	0.27	0.081	0.27
10 - 19 employees	0.099	0.30	0.093	0.29	0.10	0.31
20 - 49 employees	0.14	0.35	0.15	0.35	0.14	0.35
50 - 199 employees	0.23	0.42	0.25	0.43	0.22	0.41
200 - 499 employees	0.14	0.34	0.14	0.35	0.13	0.34
500 - 999 employees	0.080	0.27	0.076	0.26	0.083	0.28
1000 employees or more	0.14	0.35	0.14	0.34	0.15	0.35
Number of employees not provided	0.031	0.17	0.021	0.14	0.038	0.19
monthly_income	1984.3	2271.4	2285.3	3253.7	1759.4	993.5
working_hours	36.4	8.76	39.1	8.13	34.4	8.70
good health	2.07	0.78	2.00	0.76	2.12	0.79
psychological well-being	15.7	4.96	16.3	4.79	15.2	5.02
Paris area	0.13	0.33	0.13	0.33	0.13	0.34
Observations	31038		13296		17742	

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Outliers are excluded. Weights are not used.

*Legend:* "sd" is for: "standard deviation". They are not reported for dummies. Tenure is in months, and age is in years.



**TABLE A.III.**  
Raw percentages of the dependent variables and of the focal variable

	Commuting Time $T_i$		Days of sick leave absence $A_i$		Days of presenteeism $P_i$	
	mean	sd	mean	sd	mean	sd
All	24.27	20.11	3.941	10.56	2.563	9.542
Women	23.64	19.69	4.565	11.35	3.241	10.88
Men	24.89	20.49	3.340	9.700	1.911	7.998
Age $\leq$ 43	24.38	19.89	3.782	9.958	2.470	7.872
Age $>$ 43	24.15	20.36	4.124	11.22	2.670	11.18
Single	24.18	20.35	4.106	10.92	2.858	11.04
In couple	24.30	20.03	3.890	10.45	2.472	9.029
No Children	24.06	19.90	3.858	10.46	2.597	11.12
With Children	24.42	20.25	3.996	10.63	2.539	8.309
With Children under 3	25.64	21.02	4.448	11.07	2.411	6.104
at most the primary school certificate	19.53	16.37	5.209	12.94	2.640	10.17
at most the secondary school diploma	22.19	19.39	5.148	11.92	3.336	12.46
at most the certificate of professional competence	20.41	17.00	4.829	11.86	2.721	9.826
at most the baccalaureate	23.09	20.02	4.007	10.35	2.681	11.14
any higher education diploma	28.11	21.74	3.043	9.027	2.328	8.065
tenure under 219 months	24.67	20.26	3.817	10.24	2.460	8.760
tenure over 219 months	23.54	19.79	4.171	11.12	2.756	10.86
Part time job	21.69	19.38	4.556	11.82	2.875	10.93
Full time job	24.80	20.21	3.814	10.28	2.499	9.235
Permanent contract	24.45	20.12	4.148	10.83	2.646	9.837
Non-permanent contract	22.65	19.96	2.003	7.361	1.820	6.304
Public Sector	22.44	19.03	4.369	11.16	2.849	10.07
Non-public sector	24.91	20.43	3.792	10.34	2.464	9.352
Hospital sector	22.90	16.48	4.728	11.55	2.927	10.91
Non hospital sector	24.38	20.35	3.880	10.48	2.535	9.430
Public hospitals	22.62	16.58	4.622	11.52	3.057	12.05
Blue collars	21.41	16.66	4.519	11.20	2.426	9.685
Not blue collars	25.31	21.13	3.730	10.31	2.612	9.490
$<$ 50 employees	21.56	18.64	3.920	10.68	2.638	9.135
50-499 employees	25.27	20.20	4.170	10.69	2.684	9.645
500 employees or more	29.73	21.88	3.773	9.908	2.300	10.29
monthly income under 1984 euros	21.16	17.89	4.609	11.53	2.844	10.53
monthly income over 1984 euros	29.24	22.32	2.873	8.671	2.160	7.713
36 workings hours or less	22.07	18.98	4.414	11.28	2.638	9.884
more than 36 working hours	26.25	20.81	3.528	9.881	2.491	9.136
good health $\leq$ 2	24.45	20.76	7.106	14.01	5.494	16.31
good health $>$ 2	24.23	19.91	2.974	9.036	1.669	5.860
psychological well-being $\leq$ 15	25.48	20.75	4.797	11.25	4.125	12.88
psychological well-being $>$ 15	23.28	19.43	3.373	10.02	1.528	6.248
Paris area	35.77	24.94	3.158	8.861	2.479	8.285
rest of France	20.93	17.08	4.167	10.99	2.587	9.880

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Outliers are excluded ( $A_i > 90$ ). Weights are used.

*Legend:* "sd" is for: "standard deviation". They are not reported for dummies. Tenure is in months, and age is in years.

## B The respective qualities of the Poisson vs the Negative Binomial distribution

### B.1 Definitions and basic properties

Let  $Y$  be a count variable and  $X$  a vector of explanatory variables. In all the following, we make a conditional mean assumption and a distribution assumption. In the pooled regressions, the conditional mean assumption writes:

$$E(y_{i,t}|x_{i,t}) = \exp(x_{i,t}\beta) := \mu(x_{i,t}) \quad (2)$$

In the fixed-effect regressions, with  $\theta_i$  being the time-constant effect, the conditional mean assumption writes:

$$E(y_{i,t}|x_{i,t}) = \theta_i \exp(x_{i,t}\beta) := \mu(x_{i,t}) \quad (3)$$

With the vocabulary of Generalized linear models (GLM), the above regression implies that I use the log (that is the reciprocal function of the exponential) as the link function of the GLM.

A Poisson regression implies that the Poisson distribution is chosen to parameter the conditional density of  $y$  given  $x$ :

$$f(y|x) = \exp(-\mu(x)) \frac{\mu(x)^y}{y!} \text{ for } y=0, 1, \dots \quad (4)$$

A property of the Poisson distribution is the equality of the conditional variance and mean:

$$\text{var}(y|x) = E(y|x) = \mu(x) \quad (5)$$

Conversely, with  $\Gamma(x) = \int_0^\infty z^{x-1} \exp(-z) dz$  the gamma function and  $\zeta$  a positive parameter, the Negative binomial distribution assumption writes:

$$f(y_i|x_i) = \frac{\Gamma(y_i + \zeta)}{y_i! \Gamma(\zeta)} \left( \frac{\zeta}{\zeta + \mu(x)} \right)^\zeta \left( \frac{\mu(x)}{\zeta + \mu(x)} \right)^{y_i} \text{ for } y=0, 1, \dots \quad (6)$$

Conditional variance differs from the mean with the Negative binomial distribution:

$$\text{var}(y|x) = \mu(x)\left(1 + \frac{\mu(x)}{\zeta}\right) \quad (7)$$

The Negative Binomial specification is frequently used in the literature (along with the linear fixed-effect specification) as it takes better account of the overdispersed nature of the count-data under study.<sup>36</sup>

## B.2 The Poisson specification is more robust if we are only interested in the estimation of the $\beta$

However, although it makes the Negative Binomial specification a better choice when the aim is to reproduce the aggregate distribution of the dependent variable, the Poisson specification is in fact better suited to my research question. This is because I am only interested in estimating the  $\beta$  coefficients. With the Poisson specification, the estimation uses only equation (2) or (3) to get the coefficient of interest. As long as this equation is verified, the Poisson estimates of the  $\beta$  are consistent, which is not the case with a Negative Binomial specification (for which the estimates also depend on the underlying distribution).<sup>37</sup>

In the specific case of the fixed-effect specification, Wooldridge (1999b) showed that the fixed effect Poisson estimator "*is consistent for  $\beta$  under [conditional mean assumption]. Except for the conditional mean, the distribution of  $y_{i,t}$  given  $(\theta_i, x_i)$  is entirely unrestricted; in particular, there can be overdispersion or underdispersion in the latent variable model.*" (Wooldridge, 1999a).

---

<sup>36</sup>Both the number of sick leave days and days of presenteeism are overdispersed. When applying pooled regressions in Annex D, the alpha constant is significantly different from 1 (which is to be tested) for both the sick leave absence regression and the presenteeism regression.

<sup>37</sup>As I am also interested in the inference regarding the significance of those coefficients, it is also important that the standard errors are estimated correctly. This seems to be sufficiently corrected by using the robust standard errors.(Gould, 2011) The one implemented in the Stata command `xtpoisson` that I used implements the method described in Wooldridge (1999b).

### B.3 The incidental parameter problem and the use of conditional specification

The incidental parameter problem has been defined by Neyman, Scott et al. (1948) (Lancaster, 2000). In non-linear fixed-effect models where it can be considered that there is almost an infinity of individuals but the number of periods is small (i.e.  $N \gg T$ , which is the setting of the study as  $N \approx 15,000 \gg T = 2$ ), a poor estimation of the fixed-effect will bias the  $\beta$  coefficients. This requires special estimation methods to net out the  $\beta$  estimates from being influenced by the estimates of the fixed effect.

I use Stata command `xtpoisson` that performs a conditional maximum log likelihood estimation. Indeed, when conditioning on sufficient statistics (the sum of the  $y_i$ ), it can be shown that the  $\beta$  can be estimated without estimating the parameter  $\theta_i$  when the Poisson distribution is used (Hausman et al., 1984).

Such conditioning (to net out the fixed-effects) raises questions of its own when the Negative Binomial distribution is used. As noticed by Allison and Waterman (2002), it is then not truly a fixed-effect specification because fixed variables (such as gender) are still identified. This notably led Allison and Waterman (2002) to recommend using rather the unconditional negative binomial specification, hoping that the incidental parameter estimate will not be a problem, and to correct the estimated standard errors manually. As their only evidence that the incidental parameter problem will not particularly affect the Negative Binomial distribution is based on one Monte-Carlo simulation, I however still considered the Conditional Negative Binomial fixed-effect regression. Note that even in this setting, no direct method is implemented to compute the corresponding robust standard errors. They have to be estimated through (50) bootstrap replications.

## C Technical Annex: do coefficients between regressions differ statistically?

At various points, I am interested in the question of whether the difference  $d$  between two coefficients of interest for  $T_i$  in two different regression is statistically significant. In all cases, the two regressions are nested: that is regression (2) is the same as regression (1), plus some explanatory variables. I note the difference of interest:

$$d = \beta_{T_i}^{(2)} - \beta_{T_i}^{(1)}$$

I performed this significance test through a Hausman test. Note that such a test considers the reduced equation (without the added variables) to reflect the true underlying model (Clogg, Petkova and Haritou, 1995). Clogg, Petkova and Haritou (1995) presented a competing test which should be applied when the extended equation (with the added variables) is supposed to reflect the true underlying model, but this test is only valid on cross-sectional data (Yan, Aseltine Jr and Harel, 2013). Although Yan, Aseltine Jr and Harel (2013) showed how to perform this test for pooled regressions, the corresponding procedure is so far not available in Stata.

Note that these Hausman tests are performed thanks to the user-written command `rhausman` (based on 50 bootstrap replications) as the usual Hausman test is not straightly applicable to panel data with clustered and robust standard errors (Kaiser, 2015).

Table C.IV presents the tests that are performed when the test seems of interest.

TABLE C.IV.

Testing whether coefficients of interest are statistically different from each other

First equation	Second equation	$\mathbf{d}$	$\mathbf{V}(\mathbf{d})$
Table (4), column (1)	Table (4), column (2)	0.0004999***	6.798e - 09
Table (4), column (2)	Table (4), column (3)	-0.0010933***	6.442e - 08
Table (4), column (5)	Table (4), column (4)	0.0008079***	4.370e - 08
Table (4), column (6)	Table (4), column (5)	-0.0008893**	1.452e - 07
Table (5), column (1)	Table (5), column (2)	0.0000886	2.090e - 08
Table (5), column (2)	Table (5), column (3)	-0.0019781***	1.254e - 07
Table (5), column (5)	Table (5), column (4)	0.0003006*	3.196e - 08
Table (5), column (6)	Table (5), column (5)	-0.0026436***	5.678e - 07
Table (4), column (1)	Table (6), column (1)	-0.0016724	3.387e - 06
Table (4), column (2)	Table (6), column (2)	-0.0019286	3.135e - 06
Table (4), column (3)	Table (6), column (3)	-0.0007436	3.359e - 06
Table (4), column (4)	Table (6), column (4)	-0.0001862	1.424e - 06
Table (4), column (5)	Table (6), column (5)	-0.0010158	1.235e - 06
Table (4), column (6)	Table (6), column (6)	-0.0008566	6.951e - 07
Table (6), column (2)	Table (6), column (3)	0.0000917	1.787e - 07
Table (6), column (5)	Table (6), column (6)	-0.0007301***	2.714e - 08
Table (7), column (2)	Table (7), column (3)	-0.0012456	7.571e - 07
Table (7), column (5)	Table (7), column (6)	-0.0007560***	2.652e - 08
Table (5), column (1)	Table (7), column (1)	-0.0031753	4.757e - 06
Table (5), column (2)	Table (7), column (2)	-0.0029471	5.998e - 06
Table (5), column (3)	Table (7), column (3)	-0.0022147	4.778e - 06
Table (5), column (4)	Table (7), column (4)	-0.0039345**	2.901e - 06
Table (5), column (5)	Table (7), column (5)	-0.0042580**	3.376e - 06
Table (5), column (6)	Table (7), column (6)	-0.0023704**	1.349e - 06
Table (8), column (1)	Table (8), column (2)	0.0001371	2.142e - 08
Table (8), column (3)	Table (8), column (4)	0.0000788	1.199e - 07
Table (8), column (4)	Table (8), column (5)	0.0005330**	5.911e - 08
Table (9), column (1)	Table (9), column (2)	-0.0000229	2.699e - 08
Table (9), column (3)	Table (9), column (4)	0.0003843	7.216e - 07
Table (9), column (4)	Table (9), column (5)	0.0003324	2.697e - 07
Table (20), column (1)	Table (20), column (2)	0.0002059	3.415e - 08
Table (20), column (2)	Table (20), column (3)	-0.0000207	2.168e - 07
Table (20), column (5)	Table (20), column (4)	-0.0000115	4.458e - 07
Table (20), column (6)	Table (20), column (5)	-0.0007275	5.052e - 07
Table (21), column (1)	Table (21), column (2)	0.000286	6.603e - 08
Table (21), column (2)	Table (21), column (3)	-0.0014412**	4.416e - 07
Table (21), column (5)	Table (21), column (4)	-0.0000187	2.907e - 09
Table (21), column (6)	Table (21), column (5)	-0.0006949***	7.820e - 09

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Outliers are excluded. Weights are not used.

*Legend:*  $\mathbf{d}$  stands for the difference between the coefficient for  $TiTrue$  in the regression of the second column, and the same estimand in the regression of the first column.  $\mathbf{V}(\mathbf{d})$  stands for its variance according to the Hausman test.

## D Pooled regressions

TABLE D.I. Pooled regressions

	Sick leave absences $A_i$		Presenteeism $P_i$	
	(1)	(2)	(3)	(4)
	Poisson	NegBin	Poisson	NegBin
Ti	0.00136 (0.000893)	0.00230** (0.000994)	0.00304** (0.00155)	0.00250** (0.00109)
male	-0.277*** (0.0395)	-0.293*** (0.0453)	-0.308*** (0.0611)	-0.295*** (0.0563)
age	-0.0706*** (0.0165)	-0.0751*** (0.0201)	-0.0398 (0.0287)	-0.0283 (0.0210)
age2	0.000733*** (0.000191)	0.000776*** (0.000233)	0.000364 (0.000351)	0.000162 (0.000255)
In_couple	0.0194 (0.0406)	0.0180 (0.0461)	-0.0132 (0.0661)	-0.0670 (0.0551)
children	0.0889** (0.0398)	0.0674 (0.0451)	-0.0254 (0.0698)	-0.0107 (0.0526)
children_under_3	0.116** (0.0569)	0.178*** (0.0635)	-0.00920 (0.0612)	0.0445 (0.0657)
primary education	0.293*** (0.0695)	0.304*** (0.0826)	0.126 (0.116)	0.209** (0.0976)
secondary education	0.196** (0.0782)	0.277*** (0.0905)	0.197 (0.137)	0.172* (0.103)
CAP	0.368*** (0.0456)	0.408*** (0.0532)	0.152** (0.0681)	0.201*** (0.0576)
baccalaureate	0.150***	0.215***	0.0919	0.161**
Continued on next page				

Table D.I – continued from previous page

	Sick leave absences $A_i$		Presenteeism $P_i$	
	(1)	(2)	(3)	(4)
	Poisson	Negative Binomial	Poisson	Negative Binomial
	(0.0475)	(0.0544)	(0.0747)	(0.0715)
tenure	0.00120** (0.000529)	0.00231*** (0.000606)	0.00101 (0.000818)	0.00144** (0.000617)
tenure2	-0.00000240** (0.00000100)	-0.00000419*** (0.00000114)	-0.00000129 (0.00000180)	-0.00000219* (0.00000127)
part_time	-0.0886* (0.0530)	-0.119* (0.0611)	0.0435 (0.0795)	0.00456 (0.0707)
permanent	0.604*** (0.0879)	0.552*** (0.0906)	0.294*** (0.0908)	0.252*** (0.0897)
public	0.105* (0.0575)	0.153** (0.0622)	0.0506 (0.0774)	-0.00768 (0.0847)
hospital	0.134 (0.0839)	0.204** (0.0918)	-0.0914 (0.0965)	-0.177* (0.102)
public_hospital	-0.157 (0.0976)	-0.209* (0.109)	0.124 (0.130)	0.146 (0.130)
blue_collar	0.145*** (0.0469)	0.165*** (0.0545)	0.0831 (0.0643)	0.0815 (0.0577)
< 5 employees	-0.0201 (0.166)	0.0385 (0.169)	-0.0705 (0.221)	-0.127 (0.170)
5 - 9 employees	0.176 (0.159)	0.273* (0.160)	0.0249 (0.215)	0.110 (0.175)

Continued on next page



**Table D.I – continued from previous page**

	Sick leave absences $A_i$		Presenteeism $P_i$	
	(1)	(2)	(3)	(4)
	Poisson	Negative Binomial	Poisson	Negative Binomial
10 - 19 employees	0.185 (0.156)	0.299* (0.155)	0.181 (0.225)	0.204 (0.182)
20 - 49 employees	0.260* (0.153)	0.392** (0.153)	0.0871 (0.218)	0.143 (0.179)
50 - 199 employees	0.246 (0.152)	0.375** (0.152)	0.0885 (0.217)	0.0949 (0.172)
200 - 499 employees	0.243 (0.156)	0.379** (0.156)	0.0255 (0.221)	0.104 (0.178)
500 - 999 employees	0.131 (0.161)	0.206 (0.160)	-0.0379 (0.229)	0.0881 (0.188)
1000 employees or more	0.261* (0.157)	0.367** (0.157)	0.165 (0.228)	0.115 (0.184)
annee== 2016.0000	0.189*** (0.0308)	0.189*** (0.0357)	0.287*** (0.0447)	0.344*** (0.0426)
log_monthly_income	-0.184*** (0.0402)	-0.277*** (0.0603)	-0.0226 (0.0726)	-0.0686 (0.0639)
bsante	-0.506*** (0.0218)	-0.490*** (0.0237)	-0.517*** (0.0418)	-0.469*** (0.0306)
working_hours	-0.00118 (0.00263)	-0.00110 (0.00332)	0.0109** (0.00439)	0.0151*** (0.00391)
Continued on next page				

Table D.I – continued from previous page

	Sick leave absences $A_i$		Presenteeism $P_i$	
	(1)	(2)	(3)	(4)
	Poisson	Negative Binomial	Poisson	Negative Binomial
who	-0.0146*** (0.00341)	-0.0165*** (0.00384)	-0.0715*** (0.00514)	-0.0856*** (0.00416)
Constant	4.516*** (0.420)	5.041*** (0.576)	3.366*** (0.737)	3.512*** (0.619)
lnalpha constant		2.210*** (0.0151)		1.277*** (0.0240)
Observations	26584	26584	26242	26242
Pseudo $R^2$	0.09053	0.01014	0.15751	0.03503

*Source:* French Working Condition Survey 2013-2016, 18-65 year old wage earners in both 2013 and 2016 for which the focal variable (commuting time) and all the necessary controls are available.

Outliers are excluded.

## **E Pooled regressions on the additional controls**

**TABLE E.II.**  
Pooled regressions on the additional controls

Specification	Log of income		Working hours		Self-assessed health		Psychological well-being	
	(1) OLS	(2) OLS	(3) Poisson	(4) Ordered logit	(5) Poisson	(6) Ordered logit		
Ti	0.00165*** (0.000139)	0.00731*** (0.00283)	-0.000468*** (0.000102)	-0.00347*** (0.000750)	-0.000864*** (0.000128)	-0.00490*** (0.000693)		
log(monthly income)			0.0449*** (0.00544)	0.320*** (0.0391)	0.0243*** (0.00629)	0.117*** (0.0367)		
working hours			-0.00103*** (0.000287)	-0.00732*** (0.00212)	-0.00300*** (0.000344)	-0.0167*** (0.00196)		
Controls	Yes	Yes	Yes	Yes	Yes	Yes		
Observations	26764	26764	26764	26764	26764	26764		
Controls	Yes	Yes	Yes	Yes	Yes	Yes		
Fixed effects	No	No	No	No	No	No		

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Outliers are excluded. Weights are not used.

## **F The theoretical downward bias pointed out by Van Ommeren and Gutiérrez-i Puigarnau (2011).**

The dependent variables are both observed over the last 12 months, whereas the focal variable is instantaneous. If the change in the focal variable occurred one month before the interview, for instance, it might have an impact only on 1 month out of 12, whereas I calculate as if it has had an impact over 12 months. In a setting with yearly observations, Van Ommeren and Gutiérrez-i Puigarnau (2011) outlined (in their Annex B) the formal proof that it would theoretically halve their estimated coefficient compared to the coefficient of interest.<sup>38</sup> Although the general argument remains valid, such downward bias is probably less of a concern in my particular case. When the study is made on every yearly observation as in their case, such an issue affects every observation used to infer the coefficient of interest, as the change in commuting time necessarily occurred over the past 12 months and the fixed-effect regression uses them all for the identification. But in my case, the change in commuting time occurred over the three years between the two observations. Among the respondents who have changed their commuting time over the three years between the two interviews, on average only a third<sup>39</sup> has changed it over the last 12 months. The eventual downward bias would come only from them, and not from all the observations used for the inference, as was the case for Van Ommeren and Gutiérrez-i Puigarnau (2011).

---

<sup>38</sup>To avoid it, they discarded all information with a change in commuting (distance) during the year of observation. I cannot afford to do so as I do not know when the change in commuting time occurred. Note that discarding the observation or not has precisely no effect on their estimate (comparing their main estimate with the estimates of column (3) of their Appendix Table C2).

<sup>39</sup>Probably a bit more, as a few that have had a change before have also had a change later.

## G Plausible factors changing commuting time when the job and place of residence are kept constant

In the following, I try to understand what might explain a change in commuting time when the job and place of residence are constant.

I first describe what I mean by "keeping the job and place of residence constant". Change in the job was addressed in a similar way to the related literature. I used two questions. First, the interviewee was asked: "[*At the first interview,*] *you were in profession CTPROFESSC in firm CTEMPLXC. Is this still the case ?*". The interviewer was told that both profession and firm should not have changed, and that if the firm had changed while the profession had not changed, the answer should be "no".<sup>40</sup> This defines "change\_job1". Secondly, I considered the start of the contract, as provided in the 2016 wave. If it started after 2013 (or in 2013), I included it in a dummy equal to 1 in 2016. This defines "change\_job2". These two cases concerned 16 % of the (unweighted) sample (see Table G.I below). To spot changes in place of residence, I initially used the region and residence and the anonymous housing form number ("*ident\_log*"), as both are given by the survey provider. I built a dummy equal to 1 in 2016 whenever the region of residence or the *ident\_log* number changed between the two waves. This concerned roughly 2 % of the sample. The last figure seems rather low, especially when, in France, roughly one household out of 5 moves from place of residence over 4 years (as computed by Delance and Vignolles (2017) over 2009-2013). In fact, such method does only identify some changes caused by a split in the household cell or a change in the region of residence. Additional data were provided by the DARES to correctly identify changes in residence<sup>41</sup>: I now find that around 20% of people change their place of residence between the two waves.

<sup>40</sup>The French text is: "*En "V1MENQ" "V1ANEQ", vous étiez donc CTPROFESSC dans l'établissement CTEMPLXC. Est-ce toujours le cas ? (Répondre "oui" dans le cas de la même profession dans le même établissement. Si l'enquêté exerce la même profession mais dans un établissement différent, répondre " non ".)*" An "établissement" is one firm located at a given address, possibly within a broader company.

<sup>41</sup>I thank again Amélie Mauroux for this *ad hoc* data delivery.

**TABLE G.I.**  
Job and Home changes between the two waves

	weighted mean (in %)	unweighted mean (in %)
change_job1	17.8 %	13.6 %
change_job2	15.5 %	9.9 %
change_job	21.2 %	15.9 %
change_home1	2.5 %	1.4 %
change_home2	17.2 %	14.2 %
change_home	19.7 %	15.6 %

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available. Outliers are excluded. Weights are not used.

*Legend:* The variable "change\_job" spots a change either in "change\_job1" or in "change\_job2".

Second, I replicate Table 2 on the restricted sample. As I withdraw people experiencing a change in job or in place of residence, there are fewer people experiencing a "true" change in commuting time. Only 15% experience a change in commuting time of more than 10 minutes (see Table G.II) against 20% in the main sample. Similarly, the 5<sup>th</sup> and 95<sup>th</sup> percentiles are changes of -15 minutes and +20 minutes respectively <sup>42</sup> against -20 and +25 in the main sample. In the following, I try to understand what might explain changes in commuting time of more than 15 minutes in absolute terms <sup>43</sup> (in the following: "large changes") when neither the place of residence nor the job has changed.

Third and for that purpose, I look at the different explanations put forward by the various authors having used this strategy on commuting distance or time.

<sup>42</sup>It is -15 minutes and +15 minutes when weights are not used.

<sup>43</sup>This threshold of 15 minutes is chosen as it is sufficiently high for it to be unlikely that a change of more than 15 minutes is explained by imprecise answers, even in the two waves. And it is sufficiently low to concern at least 10% of the (unweighted) sample.

TABLE G.II.

Commuting time: levels and within-individual changes in the restricted sample

	all	men	women
Commuting Time $T_i$ (in min.)			
mean	24.0	24.7	23.3
sdt. dev.	19.7	19.9	19.5
Change in Commuting Time $\Delta T_i$ (in min.)			
mean	0.2	0.5	-0.2
sdt. dev.	13.3	13.5	13.2
p(5)	-15.0	-15.0	-15.0
p(10)	-10.0	-10.0	-10.0
p(50)	0.0	0.0	0.0
p(90)	7.0	10.0	5.0
p(95)	15.0	15.0	15.0
% with exactly no change	46.1 %	45.1 %	47.2 %
% with change no higher than 5 minutes	79.5 %	79.1 %	80.0 %
% with change no higher than 10 minutes	87.1 %	87.0 %	87.1 %

*Source:* French Working Condition Survey 2013-2016, 18-65 year old wage earners in both 2013 and 2016 for which the focal variable (commuting time) and main controls are available, and that have not moved or change job between the two waves. Outliers are excluded. Weights are not used.

*Lecture:* p(x) is the  $x^{th}$  percentile ; std. dev. the standard deviation.

Most of them have in mind that the firm may relocate<sup>44</sup> (Goerke and Lorenz, 2017; Gimenez-Nadal, Molina and Velilla, 2018; Künn-Nelen, 2016; Gutiérrez-i Puigarnau and van Ommeren, 2010; Munford et al., 2018).

Van Ommeren and Gutiérrez-i Puigarnau (2011) dismissed the possibility that the worker might voluntarily move to a different firm within a company, as most respondents would arguably have stated their job has changed.<sup>45</sup> Some authors, studying the impact of commuting time, also thought of changes in mode of commuting or of infrastructure<sup>46</sup> (Künn-Nelen, 2016; Roberts, Hodgson and Dolan, 2011; Munford et al., 2018).

<sup>44</sup>Thus, Goerke and Lorenz (2017) stated: "A variation in commuting distance will, thus, only occurs if a firm alters location". Or Gutiérrez-i Puigarnau and van Ommeren (2015) wrote: "controlling for job moves [makes] control for worker's voluntary workplace location changes, because it seems impossible for workers to voluntarily move to another workplace location without moving jobs (except maybe for workers who belong to top management of organisations). Munford et al. (2018) think that "[such individuals] cannot directly affect either firm/job relocation".

<sup>45</sup>As they put it: "Generally speaking a voluntary relocation to another workplace (e.g. another branch of a bank) will imply changes in colleagues, changes in job content, etc., so most of the respondents in the survey would have stated that they would have changed job." (Van Ommeren and Gutiérrez-i Puigarnau (2011), p3)

<sup>46</sup>As stated by Roberts, Hodgson and Dolan (2011), [ For a subset of the sample who do not change job, home or mode of commuting throughout the time period, ] any change in commuting time is truly exogenous, caused say by an increase in congestion".



Fourth, I investigate questions enabling plausible reasons for changes in commuting time of more than 15 minutes to be investigated (see Table G.III). First, I show that firm relocations might explain only up to 20% of the large changes in the sample. Secondly, this enables me to think of other possibilities implying voluntary changes in commuting time that were mostly overlooked by the previous literature such as jobs implying frequent changes in workplaces, or living in different places, and that are found to be both not so rare and significantly correlated with a change in commuting time. Altogether, this enables me to find a plausible explanation for up to 50% of changes in commuting time in my sample. In the end, the remaining changes (without plausible explanation) might come from changes in commuting mode (which can be endogenous to the respondent health) or from transport infrastructure (which can be thought as exogenous), or from other overlooked factors (which can be either endogenous to the respondent health or exogenous).

Finally, to take the identified time-varying factors into account, I build new variables (see in Table G.IV). They are mostly built to identify (plausible) shocks on the workplace happening in 2016, similarly to the previous core change variables (that is: `change_job1`, `change_job2`, `change_home1`).<sup>47</sup> As the kind of workplace and the fact of not living at home most of the year might have a direct impact on both absences/presenteeism and commuting, I also consider them directly.

---

<sup>47</sup>I do not add such variables identifying a change in 2016 for the variables regarding a plausible change in place of residence, due to the low number of people concerned.

**TABLE G.III.**  
Plausible explanations of the large change in commuting time

Name	Questionnaire	impact on $P( \Delta T_i  \geq 15)$	% of the $ \Delta T_i  \geq 15$
<b>all: C1 or C2 or C3 or C4</b>			<b>49 %</b>
<b>C1 - Relocation or restructuring of the firm</b>			<b>20 %</b>
C1A	<i>The address of the establishment where you work is [adress provided in wave 1]. Is that correct?</i>		
	C1A==1 : No, the firm moved	+14.0%***	5.0%
	C1A==0 : Other modalities	ref. [ 11.4%]	-
C1B	<i>In the last twelve months, has your working environment been significantly affected by a restructuring or a relocation of the [firm] where you work?</i>		
	C1B==1 : Yes	+3.1%***	17.7%
	C1B==0 : No	ref. [ 11.3%]	-
<b>C2 - In a job where the workplace varies often (in 2016 or in 2013)</b>			<b>25 %</b>
C2A	<i>Where do you spend most of your working time? If it changes often, take the last week worked.</i>		
	C2A==0 : Reference modality	ref. [ 11.1%]	-
	C2A==1 : In different firms of the company that employs you	+ 6.3%***	5.3%
	C2A==2 : In a firm to which you are sent by your employer	+ 8.9%***	2.5%
	C2A==3 : In different companies that are customers of your employer	+ 10.5%***	2.0%
	C2A==4 : On the move (ex: sales representative, truck driver...)	+ 2.7%	4.0%
	C2A==5 : Other situation	+ 3.6%**	+ 3.4%
C2A	Built variable : if C2A==0 in 2016 but not in 2013, I take the 2013 value.		
	C2B==0 : Reference modality	ref. [ 10.9%]	-
	C2B==1 : In different firms of the company that employs you	+ 5.5%***	8.9%
	C2B==2 : In a firm to which you are sent by your employer	+ 5.7%***	3.7%
	C2B==3 : In different companies that are customers of your employer	+ 11.5%***	2.8%
	C2B==4 : On the move	+ 2.3%	4.8%
	C2B==5 : Other situation	+ 2.7%*	+ 4.5%
<b>C3 - The respondent seems not to have been truly living in the same home at both waves.</b>			<b>1 %</b>
C3A	<i>Do the respondent live in his home most of the year ? At other frequencies ?</i>		
	C3A==1 : A change between 2013 and 2016.	+ 30.8%***	1.0%
	C3A==0 : No change between 2013 and 2016.	ref. [ 11.7%]	-
C3B	<i>In case the respondent is in a couple, does the partner live at her home?</i>		
	C3B==1 : A change (for respondents in a couple in both waves)	+ -3.7%* * *	0.4%
	C3B==0 : No change (Reference modality)	ref. [ 11.8%]	-
<b>C4 - A possible change in job or a firm relocation previously unnoticed.</b>			<b>17 %</b>
C4A	<i>The address of the establishment where you work is [adress provided in wave 1]. Is that correct?</i>		
	C4A==1 : No, you work in another firm.	+22.4%***	10.4%
	C4A==0 : Other modalities	ref. [ 10.9%]	-
C4B	<i>For the past year, have the people you regularly work with been the same?</i>		
	C4B==1 : No, most or all of them have changed	+5.9%***	6.2%
	C4B==0 : Other modalities	ref. [ 11.5%]	-
C4C	<i>How many other jobs of more than 1 year since the first wave?</i>		
	C4B==1 : At least one.	+17.6%***	1.9%
	C4B==0 : None.	ref. [ 11.6%]	-

Source: French Working Condition Survey 2013-2016, 18-65 year old wage earners in both 2013 and 2016 for which the focal variable (commuting time) and main controls are available, and that have not moved or change job between the two waves. Outliers are excluded. Weights are not used.

TABLE G.IV. Additional variables

<i>Variable</i>	<i>Definition</i>
<b>Core change variables</b>	
change_job1	<i>At the first interview, you were in profession CTPROFESSC in firm CTEMPLXC. Is this still the case ?</i> change_job1== 1 in 2016 if answers is no. 0 otherwise.
change_job2	The current job started in or after 2013. change_job2== 1 in 2016 if it is the case. 0 otherwise.
change_home1	The anonymous address form number has changed between 2013 and 2016. change_home1== 1 in 2016 if it is the case. 0 otherwise.
<b>Additional change variables</b>	
<i>Relocation or restructuring of the firm.</i>	
AddChange1a	<i>The address of the establishment where you work is [address provided in wave 1]. Is that correct?</i> AddChange1a==1 if in 2016 and answer is "No, the firm moved". 0 otherwise.
AddChange1b	<i>In the last twelve months, has your working environment been significantly affected by a restructuring or a relocation of the [firm] where you work?</i> AddChange1b==1 in 2016 if it is the case. 0 otherwise.
<i>In a job where the workplace varies often.</i>	
Workplace dummies	5 dummies respectively equal to 1 if most of your working time is spent : - in different firms of the company that employs you (e. g. trainer working in different firms of a company) - in a firm to which you are sent by your employer (e. g. temporary workers, service providers) - in different companies that are customers of your employer (e. g. elevator repairers) - on the move (ex: flight attendant, sales representative, truck driver...) - in "Other situation" Reference modality is : in the firm that employs you, at one or more private individuals, or at your home.
AddChange2c	Workplace question suggests that the workplace is not the same between 2013 and 2016 AddChange2c == 1 in 2016 if one of the workplace dummies is equal to 1 for the individual in 2013 or in 2016. 0 otherwise.
<i>The respondent seems not to have been truly living in the same home at both waves.</i>	
LivingHome dummy	Do the respondent live in his home most of the year ? At other frequencies ? LivingHome==1 if answer is not "most of the year" but rather "more on weekends or holidays", "more on weekdays", "a few months of the year", "more rarely". 0 otherwise.
Couple dummies	2 dummies (replacing the <i>In_couple</i> variable) respectively equal to 1 if: - in a couple with the partner living in the home. - in a couple, with the partner not living in the home.
<i>A possible change in job previously unnoticed.</i>	
AddChange4d	<i>The address of the establishment where you work is [address provided in wave 1]. Is that correct?</i> AddChange4d==1 if in 2016 and answer is "No, you work in another firm.". 0 otherwise.
AddChange4e	<i>For the past year, have the people you regularly work with been the same?</i> AddChange4e==1 if in 2016 and answer is "No, most or all of them have changed". 0 otherwise.
AddChange4f	<i>How many other jobs of more than 1 year since the first wave?</i> AddChange4f==1 if in 2016 and answer is at least 1. 0 otherwise.

## H A dummy specification for large changes in commuting time

**TABLE H.I.**  
A dummy specification for large changes in commuting time

Specification	Absenteeism $A_i$			Presenteeism $P_i$		
	(1) Fixed Effect Poisson $ \Delta Ti  >$ 5 min	(2) Fixed Effect Poisson $ \Delta Ti  >$ 20min	(3) Fixed Effect Poisson $ \Delta Ti  >$ 30min	(4) Fixed Effect Poisson $ \Delta Ti  >$ 5 min	(5) Fixed Effect Poisson $ \Delta Ti  >$ 20min	(6) Fixed Effect Poisson $ \Delta Ti  >$ 30min
Large_Change_in_Ti	0.0276 (0.0527)	0.0615 (0.0640)	0.0706 (0.0827)	0.0348 (0.0710)	0.102 (0.0902)	0.0998 (0.113)
CoreCovariables	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14206	14206	14206	17494	17494	17494

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. To make regressions fully comparable, they are also not included in the pooled regressions. Standard-errors are robust and clustered.

**TABLE H.II.**  
A dummy specification for large changes in commuting time

Specification	Absenteeism $A_i$			Presenteeism $P_i$		
	(1) Fixed Effect Poisson $ \Delta Ti  >$ 5 min	(2) Fixed Effect Poisson $ \Delta Ti  >$ 20min	(3) Fixed Effect Poisson $ \Delta Ti  >$ 30min	(4) Fixed Effect Poisson $ \Delta Ti  >$ 5 min	(5) Fixed Effect Poisson $ \Delta Ti  >$ 20min	(6) Fixed Effect Poisson $ \Delta Ti  >$ 30min
Large_Change_in_Ti	0.0266 (0.0527)	0.0430 (0.0983)	-0.0678 (0.138)	0.0350 (0.0711)	0.183 (0.125)	-0.00740 (0.135)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
ChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
AdditChangeVariables	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14206	14206	14206	17494	17494	17494

*Source:* French Working Condition Survey 2013-2016, main sample of the study: 18-65 year old wage earners in both 2013 and 2016 for whom the focal variable (commuting time) and controls are available.

*Controls:* These included personal-related variables (region dummies, a dummy for living within a couple, for having children, for having children under 3, age, and educational attainment) as well as job-related variables (tenure in quadratic form, type of contract, occupational position, firm-, industry- and sector- dummies).

*Notes:* Area dummies are not used as they prevent convergence in the fixed-effects specifications, presumably due to too little variation. To make regressions fully comparable, they are also not included in the pooled regressions. Standard-errors are robust and clustered.

# Chapitre 4

## Conclusion générale

Motivée par l'étude des politiques publiques, cette thèse pourrait rencontrer différentes audiences : chercheurs, administrateurs, auditeurs, citoyens et leurs représentants. Ce faisant, elle s'expose à un large spectre de critiques possibles quant à ses limitations, et aux approfondissements thématiques qui seraient souhaitables. Je m'efforce de les identifier, en séparant ce qui relève de l'analyse coût-bénéfices, propre aux deux premiers chapitres car spécifique à l'évaluation des politiques publiques, et ce qui relève de propositions de recherche, plus intéressées par l'éclaircissement de phénomènes comportementaux et économiques.<sup>1</sup>

### 4.1 Quels bénéfices/coûts pour les politiques publiques étudiées ?

Dans le cadre des deux premiers chapitres, les politiques publiques étudiées visent à agir sur deux objectifs quantifiables : les absences (pour raison de santé) dans un cas, le taux de chirurgie ambulatoire dans l'autre cas. Ayant déterminé des indicateurs pour ces deux objectifs, les chapitres se focalisent sur l'estimation de l'effet éventuel des politiques étudiées sur ces indicateurs. Ce faisant, ils ne

---

1. Je remercie pour leurs commentaires et suggestions Eve Caroli, Laurent Cazenave-Lacrouz, Alice Halys, Hélène Hartmann, Tomáš Jagelka et Engin Yilmaz.

répondent pas nécessairement clairement à la question finale que se posent les décideurs : cette politique est-elle souhaitable ? Et notamment : ces bénéfiques font-ils plus qu'en compenser les coûts ?

Malgré la réponse tranchée que mon co-auteur et moi-même sommes en mesure d'apporter à la question de recherche du premier article - "La mise en place d'un jour de carence n'a pas réduit les absences pour raison de santé" - , une comparaison exhaustive des bénéfiques et des coûts de la politique reste à ce jour hors de notre portée.

Le coût de la politique consiste premièrement en la perte d'assurance imposée à tous les agents potentiellement exposés à un arrêt maladie. Cette perte d'assurance s'exprime par la perte du jour de carence pour ceux qui doivent prendre un jour d'arrêt maladie, et, peut-être, par du présentéisme pour ceux qui viennent travailler alors qu'ils auraient autrement demandé à bénéficier d'un congé maladie du fait de leur état de santé. A minima, toutes choses égales par ailleurs, ceci devrait occasionner une perte de bien-être. Deuxièmement, nous montrons que la prévalence des absences de l'ordre d'une semaine à trois semaines augmente. Si cela est en soi sans incidence financière directe, du fait du rôle dual de l'État, à la fois employeur et assureur, cela pourrait cependant générer des remplacements qui, eux, sont coûteux (ou des diminution de productivité lorsque les personnes ne peuvent qu'être difficilement remplacées).

Certains bénéfiques existent même si nous montrons que la prévalence totale des absences pour raison de santé n'est pas affectée.

Premièrement, l'État, en sa qualité d'assureur, avec ou sans effet sur les absences, n'a pas à dépenser l'ensemble des jours de carence. Il est à noter que cette masse est probablement plus importante lorsqu'elle est évaluée *ex-ante* (ici : en l'absence de réaction comportementale) qu'*ex-post*. En effet, d'une part, nos résultats suggèrent une diminution de l'incidence des absences (donc moins d'absences, donc moins de jours de carence). Et d'autre part, notre objet d'étude, les absences

pour raison de santé, est insensible à des comportements de report vers d'autres type d'absences (autres types de congés, ou des accidents du travail. Voir *infra.*), contrairement aux arrêts pour maladie auxquels les jours de carence sont appliqués. Il est possible que la diminution des arrêts pour maladie, soumis au jour de carence, soit plus importante que ce que nous obtenons pour les absences pour raison de santé.

Deuxièmement, en diminuant la prévalence des absences courtes au détriment des absences plus longues, la politique peut avoir un impact, positif ou négatif (voir *supra*), sur la productivité. Cet impact peut être éventuellement positif si les absences courtes ne peuvent pas être remplacées, ou si elles ont plus généralement des impacts négatifs sur l'environnement de travail.

Troisièmement, mais peut-être de façon plus subsidiaire, en présence d'effets (éventuels) de report vers des motifs d'arrêt tels des jours de vacances, ou des jours de R.T.T., (qui auraient été sinon utilisés pour partir en congé), la politique pourrait conduire à augmenter, marginalement, la durée de travail.

A l'issue du second article, il ne peut pas y avoir non plus de réponse catégorique à une comparaison entre les coûts et les bénéfices. Certes, à partir du moment où les politiques ont des effets, elles engendrent un bénéfice potentiellement important pour le système de santé en général, à des coûts supplémentaires *a priori* faibles. Mais la *répartition* de ce bénéfice varie avec l'instauration de tarifs uniques, et dépend de plus de leur fixation spécifique. De telles analyses bénéfices-coûts s'avèrent ainsi spécifiques à chacune des deux politiques étudiées (mais sans indépendance), et parfois à chacun des acteurs impliqués, voire à chaque GHM, et restent incertaines. Uniquement en ce qui concerne la Mise Sous Accord Préalable, même si ses bénéfices associés sont incertains, ils paraissent *a priori* plus élevés que ses coûts.

Je commence par l'analyse qui est commune au deux politiques, en considérant le système de santé en général. La chirurgie ambulatoire permet déjà de diminuer



les coûts de prise en charge à l'hôpital et d'améliorer la qualité des soins (par une diminution des infections nosocomiales notamment). S'il n'y a pas de consensus dans la littérature en ce qui concerne l'ampleur des autres bénéfices et coûts additionnels (par exemple, la chirurgie ambulatoire peut entraîner une hausse du coût de prise en charge par la médecine de ville, surcoût rarement pris en compte dans les études - voir ANAP et HAS (2012)), les décideurs publics trouvent dans ces deux avantages des raisons suffisantes d'encourager dans la durée un virage ambulatoire. Ainsi, dès lors que les politiques contribuent bien à développer la chirurgie ambulatoire, on peut raisonnablement penser qu'elles devraient occasionner un bénéfice potentiellement important.

Or, face à ce bénéfice, les deux politiques ne semblent pas créer de coûts directs importants pour le système de santé.

Cela est clair pour la politique du tarif unique, qui n'est *in fine* qu'une modulation des tarifs relatifs de différents groupes de malades, alors qu'une enveloppe globale (la composante des dépenses relatives aux établissements de santé de l'ONDAM, l'Objectif National des Dépenses d'Assurance Maladie) encadre les dépenses hospitalières d'une année sur l'autre. Par construction, la mise en place de tarifs uniques ne peut donc pas créer par elle-même d'augmentation des transferts envers l'ensemble des hôpitaux. En revanche, on peut noter que ces tarifs uniques font des gagnants et des perdants parmi les hôpitaux (respectivement ceux qui, au sein de leur secteur, réalisaient déjà beaucoup de chirurgie ambulatoire et ceux qui n'en réalisaient pas beaucoup, voire ne pourront pas en réaliser beaucoup du fait de la structure de leur patientèle), et ce déséquilibre peut à son tour générer des coûts (si cela génère ou amplifie des déficits d'établissements que l'Etat n'est pas prêt à laisser faire faillite).

S'agissant des coûts liés à la Mise sous Accord Préalable, on retrouve notamment le surcoût administratif lié à la mise en place en pratique d'une mise sous accord préalable d'un hôpital : à savoir, dans ces situations actuellement rares, l'emploi de personnels de l'Assurance Maladie pour analyser si certaines opérations pré-

vues par les hôpitaux en hospitalisation complète ne pourraient pas être menées en chirurgie ambulatoire. Cela semble cependant négligeable face aux effets positifs évoqués plus haut.

Les deux politiques semblent donc engendrer un bénéfice probablement important pour le système de santé pris dans sa globalité, à des coûts qui semblent faibles.

Mais ce bénéfice se répartit entre les acteurs<sup>2</sup> de façon très variable.

Cela est apparent dans le cas de la politique de la Mise sous accord préalable, même si la répartition éventuelle dépend de l'existence de tarifs uniques. En l'absence de tarif unique, c'est l'Assurance Maladie qui bénéficie directement des économies liées à des opérations réalisées en chirurgie ambulatoire plutôt qu'en hospitalisation complète. En présence de tarif unique, cette substitution semble neutre financièrement pour l'Assurance Maladie, mais bénéficie entièrement aux hôpitaux (si elle n'a pas été anticipée dans la fixation du sous-objectif de l'ONDAM qui les concerne). Même dans ce dernier cas, l'Assurance Maladie y gagne cependant à terme, puisqu'une augmentation du pourcentage d'opération en chirurgie ambulatoire peut dynamiquement rétroagir sur le niveau du tarif unique (en le baissant) dans les années suivantes - s'il est même en partie lié à la moyenne des coûts sur le GHM.

Estimer qui obtient ce bénéfice est singulièrement plus complexe dans le cas de la politique du tarif unique.

Trois cas particuliers peuvent être considérés. Premièrement, si le tarif unique est placé au tarif de la chirurgie ambulatoire : alors, on peut avoir de prime abord l'impression que l'Assurance Maladie réalise de façon opportuniste des économies sur les séjours encore en hospitalisation complète, tarifés au prix de la chirurgie ambulatoire, et sur les séjours en ambulatoire qui auraient été sans la mesure réalisés en hospitalisation complète. Deuxièmement, si le tarif unique est placé au tarif des séjours en hospitalisation complète, alors on peut avoir de prime abord l'impression que ce sont les hôpitaux qui y gagnent : ils gagnent de façon oppor-

---

2. On ne considérera ici que les hôpitaux et l'Assurance Maladie.

tuniste sur l'ensemble des séjours qu'ils réalisaient en chirurgie ambulatoire, et sur les séjours en ambulatoire qui auraient été sans la mesure réalisés en hospitalisation complète. Troisièmement, si le tarif unique est placé à la moyenne des coûts des séjours en hospitalisation complète (de niveau de sévérité 1) et en chirurgie ambulatoire, alors les effets opportunistes se compensent. Les hôpitaux ne gagnent que sur les séjours en ambulatoire qui auraient été sans la mesure réalisés en hospitalisation complète. Enfin, même dans ces trois situations simples, il faut aussi ajouter que les tarifs sont ensuite encore systématiquement réajustés, à la hausse ou à la baisse, pour que l'ensemble des dépenses hospitalières respecte une enveloppe globale, à savoir le sous-objectif de l'Ondam précédemment évoqué.

Les premiers tarifs uniques en 2009<sup>3</sup> consistaient en une moyenne des tarifs de chirurgie ambulatoire et en hospitalisation complète (de niveau de sévérité 1), moyenne pondérée par le pourcentage respectif de séjours dans le secteur (public ou privé) où la proportion d'ambulatoire était la plus élevée. Cela correspondait plutôt à un cas légèrement défavorable aux hôpitaux (considérés de façon agrégés), entre le premier et le troisième cas ci-dessus. Ainsi, en 2010, l'avantage tarifaire accordé aux GHM ambulatoires incités a rapporté moins aux hôpitaux que la diminution du tarif des GHMs de niveau de sévérité 1 correspondant (Bras, Vieilleribiere et Lesteven, 2012) (-25 millions d'euros dans le secteur public, et -11 millions d'euros dans le secteur privé).

Paradoxalement, on ne devrait cependant pas en conclure aussitôt que l'Assurance Maladie est le bénéficiaire net de la politique. Du fait de l'ajustement réalisé pour respecter l'Ondam, ces baisses de transferts envers les hôpitaux traitant les GHMs avec incitations ont théoriquement été compensés par des hausses de transferts vers les hôpitaux traitant plus les GHMs sans incitation. Cette baisse de transferts est donc un enjeu de répartition des ressources entre hôpitaux (ceux qui traitent les GHM incités, et les autres), avant d'être un enjeu de répartition entre l'Assurance Maladie et les hôpitaux pris de façon agrégés.

---

3. Pour la campagne de tarifs 2014 (non étudiée), le choix a été fait d'avoir des tarifs uniques « plutôt bas » sur les GHM où la chirurgie ambulatoire est très développée (supérieure à 50 %) et « plutôt haut » sur les GHM où la chirurgie ambulatoire est peu développée (inférieure à 10 %) et neutre sur les GHM intermédiaires (Bert et al., 2014).

Ces enjeux de répartition entre hôpitaux ne doivent cependant pas être négligés. Déjà, derrière ces distorsions techniques, on retrouve des activités de soin qui peuvent être comparativement freinées. Par exemple Bras, Vieilleribiere et Lesteven (2012) notent que les séances de chimiothérapie ou de dialyse, et l'opération de la cataracte, bénéficient ainsi d'avantages tarifaires importants alors que, dans le secteur public, la gynécologie est particulièrement pénalisée. La baisse des transferts évoquée précédemment aurait ainsi pu être utilisée pour cesser le désavantage tarifaire subi par la gynécologie. De plus, derrière les sommes agrégées évoquées *supra* pourraient se dissimuler des situations très contrastées, par exemple entre hôpitaux qui réalisaient déjà beaucoup de chirurgie ambulatoire, et hôpitaux dans un environnement où cela est intrinsèquement difficile ; et hôpitaux qui ont un *case-mix* de fait très orientés vers les GHM très concernés par le virage ambulatoire, et d'autre où cela est moins le cas. Les analyses correspondantes sont cependant au delà du cadre de notre travail.

Pour être exhaustif, et de façon non négligeable, l'Assurance Maladie peut aussi espérer bénéficier dynamiquement d'une augmentation de la chirurgie ambulatoire et du tarif unique par le biais de baisses futures du tarif unique, ce qui est cependant à nouveau difficile d'estimer.

## 4.2 Quelles nouvelles question de recherches ?

Les résultats des chapitres précédents soulèvent à leur tour de nombreuses questions de recherche intéressantes, dont l'étude a dû être laissée hors du cadre de la thèse.

Concernant le premier chapitre de la thèse, une question majeure qui reste sans réponse à ce stade est évidemment le mécanisme (ou les mécanismes) par lequel les absences de 1 semaine à 3 semaines augmentent. Comprendre quel est ce mécanisme est fondamental pour qui veut diminuer les absences de façon efficace. En combinaison avec une meilleure compréhension des déterminants de l'hétérogé-

néité des réactions, cela permettrait probablement de rendre possible des progrès dans l'étude théorique des systèmes d'indemnisation optimaux des arrêts maladies.

La désirabilité de l'objectif même de telles politiques pourrait aussi être plus étayée ; par exemple, peut-être qu'une partie des employés qui ont renoncé à des absences de courte durée était effectivement malade - c'est un pré-supposé de l'une des hypothèses que nous avançons pour expliquer l'augmentation des absences de plus longue durée. Autrement dit, il serait intéressant de déterminer à quel point les renoncements aux absences de courte durée visées par la politique impliquent du présentéisme.

De même, et en particulier dans le contexte sanitaire et économique actuel dans lequel cette thèse se clôt, des études montrent que les dispositifs d'indemnisation des arrêts maladies peuvent affecter la propagation d'épidémies (cf. Marie et Vall Castelló (2020), Pichler et Ziebarth (2017), Stearns et White (2018)). En l'espèce, il serait intéressant d'estimer dans quelle mesure le jour de carence est aussi susceptible de favoriser la diffusion d'épidémies telle celle du covid-19.

Par ailleurs, notre étude se concentre aussi sur un indicateur disponible : les absences pour raison de santé - sans mention de la situation administrative pendant cette absence. Il serait aussi intéressant de mener une étude similaire avec l'indicateur directement ciblé par la politique, les arrêts maladies. Par exemple, ces derniers pourraient diminuer du fait de la mesure, quand bien même nous n'obtenons pas d'effet sur les absences pour raison de santé. Cela pourrait être le cas par exemple s'il existe des effets de report où des absences qui auraient été déclarées comme des arrêts maladies (en l'absence de la mesure) seraient déclarées autrement du fait de la mesure (par exemple comme des accidents et maladies du travail, comme cela a pu être observé par Marie et Vall Castelló (2020) dans un contexte similaire). Des études traitant principalement de l'effet sur les arrêts maladies devraient donc être idéalement menées de façon à pouvoir aussi identifier ces effets de report.

Enfin, en supprimant la nécessité du trajet pendulaire, en rendant potentiellement

la journée de travail plus flexible et plus adaptable, et en supprimant les risques de transmission d'épidémie sur le lieu du travail, le développement du télétravail pourrait altérer les éventuels équilibres obtenus dans des études lors desquelles le télétravail restait relativement marginal.

Concernant le deuxième chapitre de cette thèse, des éclaircissements sur la possible hétérogénéité des effets des mesures suivant les caractéristiques des hôpitaux (au-delà de la simple comparaison secteur public / secteur privé) seraient intéressants. Cependant, ceux-ci s'accommodent mal avec la méthode de construction des contrôles synthétiques.

Il serait peut-être pertinent de s'intéresser de façon descriptive aux disparités d'évolution de la pratique de la chirurgie ambulatoire sur la période, suivant les caractéristiques des hôpitaux. De telles études permettraient d'identifier si certaines caractéristiques (d'hôpitaux ou de régions) expliquent un retard systématique sur des pathologies pour lesquelles la chirurgie ambulatoire pourrait être la norme, et ce malgré les politiques incitatives ; des structures avec ces caractéristiques pourraient en effet subir à terme des pertes financières importantes à mesure que le poids de la chirurgie ambulatoire augmentera dans la pondération des tarifs uniques, et il conviendrait dès lors de les identifier au plus tôt afin d'éviter la mise en place subséquente de mesures plus drastiques, ou d'avoir à les renflouer).

Extrapolant encore au-delà de la politique étudiée, l'intérêt pour la chirurgie ambulatoire s'accorde avec un discours politique sur une volonté de décroïsonner l'hôpital, en incitant ce dernier à envisager la suite du parcours de soin du patient. Cette politique trouverait peut-être une continuation logique dans une tarification à l'épisode de soin ou par la mise en place de réseaux de santé (à l'instar des *Accountable Care Organizations* américaines, voir Mousquès et Lenormand (2017)), nécessitant cependant une réflexion sur leur pertinence et les modalités de leurs éventuelles mises en place.

Concernant enfin le troisième chapitre de la thèse, pour obtenir une inférence

causale la plus convaincante possible, fondée uniquement sur des modifications de temps de trajet exogènes, car causées par des déménagements d'entreprise, il serait évidemment souhaitable de combiner données administratives (si elles existent) sur les trois variables que sont les déménagements d'entreprise, les arrêts maladies pris, et les temps de déplacement domicile travail<sup>4</sup> - même si la dimension "présentéisme" serait toujours manquante.

De façon plus immédiate, il serait aussi intéressant de considérer à part, et d'approfondir, le résultat quant à l'effet du temps de trajet domicile-travail sur la santé et le bien-être mental des femmes. De tels travaux pourraient viser à explorer les hétérogénéités éventuelles de ce résultat afin de comprendre ce qui explique cet effet différencié entre femmes et hommes. Un tel approfondissement n'a certes pas été réalisé parce qu'il ne s'agissait pas de l'objet du chapitre de la thèse, mais cela pourrait peut-être donner matière à un nouvel article en soi.

---

4. Ces derniers ne seront pas disponibles dans aucune base administrative ; il est en revanche peut-être possible de calculer une distance et un temps de déplacement moyen (en voiture ou en transport en commun) entre l'adresse du domicile et l'adresse de l'employeur.

## Bibliographie

- ANAP, et HAS.** 2012. “Ensemble pour le développement de la chirurgie ambulatoire, socle de connaissances.”
- Bert, Thierry, Christine D’Autume, Stéphanie Dupays, Pierre Hausswalt, Julie Perlberg, et Morgane Weill.** 2014. “Perspectives du développement de la chirurgie ambulatoire en France.” IGF-IGAS.
- Bras, Pierre-Louis, Jean-Luc Vieilleribiere, et Pierre Lesteven.** 2012. “Évaluation de la tarification des soins hospitaliers et des actes médicaux.” IGAS.
- Marie, Olivier, et Judit Vall Castelló.** 2020. “If Sick-Leave becomes More Costly, Will I go back to Work? Could it be too soon?”
- Mousquès, Julien, et Marie-Camille Lenormand.** 2017. “L’expérience américaine des Accountable care organizations: des enseignements pour la France.” *Irdes QES*, 227: 7p.
- Pichler, Stefan, et Nicolas R Ziebarth.** 2017. “The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior.” *Journal of Public Economics*, 156: 14–33.
- Stearns, Jenna, et Corey White.** 2018. “Can paid sick leave mandates reduce leave-taking?” *Labour Economics*, 51: 227–246.





# Table des matières

<b>Remerciements</b>	<b>1</b>
<b>Sommaire</b>	<b>5</b>
<b>Introduction Générale</b>	<b>9</b>
Les questions . . . . .	10
Les méthodes . . . . .	12
Recommandations de politiques publiques . . . . .	17
Bibliographie . . . . .	27
<b>1 Les effets d'un jour de carence pour arrêt maladie sur les absences pour raison de santé dans la fonction publique de l'État française</b>	<b>31</b>
<i>Lower sick leave cover, fewer health-related work absences ?</i>	
1.1 <i>Introduction</i> . . . . .	34
1.2 <i>Institutional framework, data, and descriptive statistics</i> . . . . .	37
1.2.1 <i>The 1st January 2012 introduction and the 1st January 2014 repeal of the one-day waiting period in the French civil service</i>	37
1.2.2 <i>Data : the French Labour Force Survey</i> . . . . .	39
1.2.3 <i>Descriptive statistics</i> . . . . .	43
1.3 <i>Empirical strategy</i> . . . . .	52
1.4 <i>Results</i> . . . . .	54
1.4.1 <i>Treatment effects for spells of different durations</i> . . . . .	54
1.4.2 <i>Heterogeneous effects</i> . . . . .	58

1.4.3	<i>Switch on and switch off</i>	60
1.5	<i>Robustness tests</i>	61
1.5.1	<i>Placebo test</i>	61
1.5.2	<i>Alternative control group</i>	62
1.5.3	<i>Covariates and fixed effects</i>	63
1.5.4	<i>Weights</i>	64
1.5.5	<i>Nonlinear specifications</i>	65
1.6	<i>Conclusion</i>	67
1.7	<i>References</i>	69
1.8	<i>Appendices</i>	73
1.8.1	<i>Questionnaire on sick leave from the French Labour Force Survey</i>	73
1.8.2	<i>Prevalence of sick leave spells over time and sector, by duration category</i>	74
1.8.3	<i>Determinants of absenteeism</i>	76
1.8.4	<i>Treatment effects for the 8 quarters of implementation of the policy</i>	77
<b>2</b>	<b>Incitations par les prix, monitoring et pratiques médicales des hôpitaux</b>	<b>79</b>
	<i>Can price incentives or the threat of monitoring induce hospitals to change their medical practices? Evidence from two French policies.</i>	
2.1	<i>Introduction</i>	82
2.2	<i>The single tariff policy and the prior approval policy</i>	87
2.2.1	<i>The practice of same-day surgery in France</i>	87
2.2.2	<i>The single tariff policy</i>	88
2.2.3	<i>The prior approval policy</i>	89
2.3	<i>Data and the studied treatments</i>	92

---

2.3.1	<i>Data</i>	92
2.3.2	<i>The two treatments</i>	94
2.4	<i>The Empirical strategy</i>	96
2.4.1	<i>Difficulties</i>	96
2.4.2	<i>The Methodology</i>	97
2.4.3	<i>Creation of a synthetic control rate of same-day surgery</i>	98
2.4.4	<i>Inference with a synthetic control estimator</i>	101
2.4.5	<i>Further restrictions</i>	104
2.5	<i>Results</i>	106
2.5.1	<i>Rates of same-day surgery for DRGs with the single tariff policy</i>	106
2.5.2	<i>Rates of same-day surgery for DRGs starting with only the prior approval policy</i>	113
2.5.3	<i>Magnitude of the effects</i>	119
2.5.4	<i>Impacts on other variables</i>	123
2.6	<i>Conclusion</i>	128
2.7	<i>References</i>	132
2.8	<i>Glossary</i>	137
2.9	<i>Appendices</i>	143
2.9.1	<i>Characteristics of the data base</i>	143
2.9.2	<i>Graphs for the DRGs incentivized starting from 2012</i>	145
2.9.3	<i>DRGs for which no reasonable synthetic control was found.</i>	150
2.9.4	<i>Results tables when the data-driven matrix <math>V</math> is used</i>	158
2.9.5	<i>Result tables when a weight of 50 to the two control variables</i>	163
2.9.6	<i>Graphs associated with some significant results regarding complementary variables</i>	168
2.9.7	<i>The higher the distance to the mid-term prospect, the higher the effect : suggestive evidence.</i>	170

**3 Temps de transport domicile-travail, absentéisme et présen-  
téisme en France** **173**

*Does commuting time affect absenteeism and presenteeism in France ?*

3.1	<i>Introduction</i>	176
3.2	<i>Data and variables</i>	180
3.3	<i>Empirical strategy</i>	184
3.4	<i>Results</i>	187
3.4.1	<i>Descriptive pooled regressions</i>	187
3.4.2	<i>Simple fixed-effect regressions</i>	192
3.4.3	<i>Results when identification is improved</i>	196
3.4.4	<i>Dealing with eventual data issues</i>	200
3.4.5	<i>Heterogeneity</i>	205
3.4.6	<i>Dealing with weights</i>	214
3.4.7	<i>Testing another focal variable</i>	216
3.5	<i>The intermediation of health in the studied relationships</i>	219
3.5.1	<i>Psychological well-being may intermediate most of the studied relationships</i>	219
3.5.2	<i>Does commuting time have a direct health effect ?</i>	223
3.6	<i>Conclusion</i>	227
3.7	<i>References</i>	230
3.8	<i>Appendices</i>	235
3.8.1	<i>Variables and sample composition</i>	235
3.8.2	<i>The respective qualities of the Poisson vs the Negative Binomial distribution</i>	239
3.8.3	<i>Technical Annex : do coefficients between regressions differ statistically ?</i>	242
3.8.4	<i>Pooled regressions</i>	244
3.8.5	<i>Pooled regressions on the additional controls</i>	248
3.8.6	<i>The theoretical downward bias pointed out by Van Ommeren and Gutiérrez-i Puigarnau (2011).</i>	250

3.8.7	<i>Plausible factors changing commuting time when the job and place of residence are kept constant . . . . .</i>	251
3.8.8	<i>A dummy specification for large changes in commuting time</i>	257
<b>4</b>	<b>Conclusion générale</b>	<b>259</b>
4.1	Quels bénéfices/coûts pour les politiques publiques étudiées? . . . .	259
4.2	Quelles nouvelles question de recherches? . . . . .	265
4.3	Bibliographie . . . . .	269
	<b>Table des matières</b>	<b>271</b>







## RÉSUMÉ

---

La thèse traite de trois sujets empiriques en économie de la santé et du travail intéressant les politiques publiques.

Le premier article analyse l'effet sur les absences pour raison de santé de la mise en place puis de la suppression d'un jour de carence, expérimenté dans la fonction publique française en 2012 et en 2013. Il y a moins d'absences courtes et plus d'absences longues, mais la prévalence totale des absences reste inchangée.

Le second article évalue l'influence des incitations financières sur les décisions médicales des hôpitaux, par le biais de l'évaluation d'une politique tarifaire française qui a pour objectif d'augmenter la proportion de séjours en chirurgie durant seulement 1 jour. Cette politique était concomitante d'une politique de mise sous accord préalable, au même objectif, mais portant sur un champ légèrement différent. Des effets sont surtout obtenus au lancement des deux politiques, en particulier dans le secteur public.

Le troisième article analyse les possibles impacts du temps de trajet domicile-travail sur les arrêts maladies et le présentéisme des salariés français. Je ne met pas en évidence de tels impacts, mais montre que le temps de trajet domicile-travail peut affecter la santé auto-évaluée et le bien-être mental des femmes.

## MOTS CLÉS

---

jour de carence, absentéisme, présentéisme, chirurgie ambulatoire, trajet domicile travail, politiques publiques

## ABSTRACT

---

The thesis addresses three empirical topics in health and labour economics of interest to public policies.

The first paper analyses the effect on health-related absences of the introduction and then removal of a one day waiting-period experienced in the French civil service in 2012 and 2013. There are fewer short-term absences and more long-term absences but the total prevalence of health-related absences remains unchanged.

The second article evaluates the influence of financial incentives on hospitals' medical decisions, through the evaluation of a French pricing policy that aims to increase the proportion of surgical stays of only 1 day. This policy was concomitant with a policy of prior agreement, with the same aim, but with a slightly different scope. Effects are mainly obtained at the launch of the two policies, especially in the public sector.

The third article analyses the possible impacts of commuting time on sick leave and presenteeism among French employees. I do not find evidence of such impacts, but show that commuting time can affect women's self-assessed health and psychological well-being.

## KEYWORDS

---

waiting period, absenteeism, presenteeism, same-day surgery, commuting, public policies