

Discontinuous System of Allowances: the Response of Prosocial Health-Care Professionals*

Helena M Hernández-Pizarro
Universitat Pompeu Fabra
CRES and TecnoCampus

Catia Nicodemo
University of Oxford
CHSEO

Guillem López Casasnovas
Universitat Pompeu Fabra
CRES

July 28, 2020

Abstract

This paper examines the unintended strategic effects of non-linear incentives in public policies. A system of health care subsidies structured into discrete intervals may lead to strategic behaviour. We provide new evidence on this issue, focusing on a case where strategic actions are taken by healthcare providers (HCPs). We show that HCPs adjust the score of claimants in long-term care needs assessments, allowing claimants to access higher levels of care benefits. This adjustment does not bring any pecuniary gain for HCPs. In a quasi-natural experimental setting –the implementation of Spanish long-term care (LTC)– we show that the distribution of the claimants’ needs includes kinks around the thresholds set for the LTC system. These kinks reveal that healthcare providers adopt prosocial behaviour, helping claimants jump to a higher level of benefits without discriminating by health status, residence, or gender. By developing a new non-parametric estimator, we prove that these adjustments increase the public LTC spending, which is equivalent to moving thresholds to the left. The additional cost per adjusted claimant per annum is approximately 1000 euro on average. Finally, we also explore empirically whether the number of thresholds could affect such behaviour, by exploiting a reform that eliminated some thresholds. We find that the lower number of thresholds is associated with an increase of adjustment around the thresholds that remain.

Keywords: Long-Term Care, Non-Linear Incentives, Needs Assessment, Prosocial Behaviour

JEL classification: D63, D82, D61, H510, I380

*Previously circulated under the title “Against the Discontinuity Trap: the Case of Long-Term in Spain”. This paper benefited from helpful conversations and suggestions from Jose Apestegua, Manuel Arellano, Johannes Abeler, David Cox, Vincent Crawford, Gabrielle Fack, Jaume Garcia, Beatriz González López-Valcárcel, Hector Pifarré, Sergi Jimenez and Pau Olivella. We would also like to thank all participants at CRES- UPF Seminar, University of Oxford Seminar, XXXIV AES Conference, UB-LTC Workshop and 10th World Congress of Health Economics (iHEA and ECHE, Dublin 2014) for helpful comments and discussion. We are grateful to the Catalan Department of Social Welfare and Family (ICASS) for help with data and policy understanding. We acknowledge financial support from the AGAUR (2017- SGR 1059), the Agencia Estatal de Investigación (AEI) (project ECO2012-37496) and Fondo Europeo de Desarrollo Regional (FEDER) (project ECO2017-88609-R). Errors remain our own.

1 Introduction

Prosocial theory, altruism, intrinsic motivation, and other phenomena, have been used to demonstrate that there are non-pecuniary aspects of work, jobs, and processes within organizations that can affect worker behaviour. Economists have emphasised that individuals often care for the well-being of others (Becker (1974); Samuelson (1993)). It has been widely documented that employees in some settings, such as education, social services, healthcare, etc. are more inclined to respond to non-market incentives.¹ In the healthcare sector the utility function of physicians is usually modelled not only on their salary, but also on patients' health status and/or patient utility (Ellis et al. (1986); Evans (1984); Feldstein (1970); Siciliani and Verzulli (2009)). Additionally, other factors may affect physicians' utility, such as intellectual stimulation, reputation and/or the intrinsic motivation from doing a good job (Fehr and Camerer (2007); Fehr and Falk (1969); Frey and Oberholzer-Gee (1997); Frey and Jegen (2001), Charness and Haruvy (1999)). In this paper, we provide empirical evidence of how prosocial motivation affects healthcare providers' (physicians, nurses, physiotherapists and social workers, henceforth HCPs) behaviour in classifying the elderly into grades of Long-Term Care (LTC), based on the severity of their care requirements and which are used to give access to benefits. LTC is required by individuals who have lost the ability to independently perform activities of daily living (ADL).² In 2007 the Spanish government implemented a universal LTC system, which classifies benefits claimants into six grades (with appropriate thresholds) based on the loss of autonomy suffered. All grades entitle individuals to some subsidised care, via vouchers or in-kind care services, with more care provided to those at a higher grade. An official scale was designed to determine the level of LTC needs. The scale ranges from 0 to 100, with patients rated 100 those with the maximum level of LTC needs. The score can be any value between 0 and 100, but benefits were defined for six grades. Consequently, the distribution of LTC needs, measured using the official scale, tends to kink around the scores that distinguish between levels of benefits, see Figure 1. We claim that these notches represent strategic behaviour of HCPs. HCPs appear to adjust the scores of claimants when the assessed level of needs (score) is just below a threshold (25, 40, 50, 65, 75 and 80), allowing claimants to access the benefits for the next grade up.

Given this hypothesis, this paper has two objectives: firstly, we document that HCPs behave strategically under a non-linear classification system and, secondly, we estimate the economic con-

¹Le Grand (2003) and Perry et al. (2010) survey the empirical literature on workers' motivation in public sector. See Francois and Vlassopoulos (2008) for a review of theoretical research on the motivation of public workers.

²Activities of daily living (ADL) include basic activities, such as bathing, and instrumental activities, such as cooking.

sequences of this behaviour in terms of public spending on LTC in Catalonia, the North-East region of Spain, using a novel administrative dataset from social services authorities (ICASS).

We show that prosocial motivation causes HCPs to help claimants become eligible for higher level of benefits, aiming to increase the claimants' comfort and well-being. We report robust evidence to prove that the distribution of the observed scores is produced by the prosocial motivation of the HCPs. We first analyse the characteristics of the claimants around each cut-off, finding no significant differences in the observable factors of individuals whose scores are just above and just below the thresholds. We also provide evidence that the adjusting behaviour occurs in other Spanish regions, and that there are no differences in the score distributions observed across local authorities areas.³ Three additional factors of the assessment process need to be stressed to support the prosocial hypothesis. First, claimants have no prior relationship with the examiners, as they meet only once when the assessment takes place. This feature eliminates other factors that could be responsible for the HCPs adjustment, such as competition with other doctors and/or reputational/satisfaction issues amongst patients. Second, the obtained score is not disclosed by the HCP to claimants during the assessment. The score is reported to the central social services' authorities, who inform the claimants only the grade assigned. This reduces the possibility of bribery, reputational or other coercive motives (for more discussion on this see Andreoni (1990)). Third, when a policy reform eliminated intermediate grades in 2012 (reducing the number of categories of benefits from six to three), we observe that the adjusting behaviour around the pre-reform cut-off points disappeared, but grew stronger around the remaining cut-offs. Furthermore, we conclude that neither altruism nor personal satisfaction can cause the strategic behaviour observed.

We first model the HCP's adjusting decision under a discontinuous scheme of benefits. The strategic actions of HCPs around the local cut-off is equivalent to a left shift of the threshold, with an increase in benefits allocated at a cost to the public purse. Second, to measure the public financial cost for the long-term care system, we develop a new non-parametric econometric model to estimate the counterfactual distribution of LTC needs (i.e. the shape of the distribution in the absence of adjustments) using data of Spanish LTC claimants between 2007 and 2010. This method allows us to recover the real distribution.⁴ We also exploit the previously mentioned 2012 reform, that reduced the number of grades from six to three, to explore the effects of the number of thresholds on the induced behaviour. We found that the reduction of thresholds increases the difference between

³Using data from the Basque Country we find that the distribution of scores follows the same pattern with kinks in other Spanish regions.

⁴It is not possible to use kink regression models in this case as the decision to be above or below to the cut-off is not causal.

the benefits obtained in each grade, and could increase the incentive of HCPs to adjust. Yet, the reduction of thresholds also reduces the total number of claimants where adjustment is likely to be considered (i.e. whose needs-scores are close to thresholds). Comparing the scenarios pre and post the 2012 reform, we observe that the percentage of adjusted claimants remains relatively constant (around 5%). However, we find that the kinks just above each remaining threshold are steeper and larger, as the share of adjusted claimants at these points of the distribution substantially increases.

This paper contributes to the literature on prosocial motivation. Although theoretical models in health economics have identified this factor for decades, empirical economic works attempting to measure it are in their infancy. Laboratory experiments have been used to identify prosocial motivation (see Levitt and List (2007) for an overview), but real-world experiments have been less well utilised. In the health-care literature the strategic behaviour of health-care professionals has usually been studied in settings with monetary incentives. Einav and Mahoney (2018) find a disproportionately high number of patients are discharged from long-term care hospitals after a duration threshold that substantially increases the Medicare reimbursement payment. This delayed discharge does not impact on patient mortality. Similarly, when Diagnosis Related Groups (DRG) are used to distribute funds, patients whose health conditions are just below thresholds can be upgraded into the more severe category of illness (Jürges and Köberlein (2015)).⁵ When incentives are linked to quality indicators, there is also evidence of strategic behaviour: Maxwell et al. (2014) identified an abrupt increase in mortality immediately after 30 days of an operation. Authors claim that pressure placed on physicians to meet a quality target (mortality within 30 days of an operation) could explain the abrupt increase. Few papers have measured the degree of altruism amongst the medical profession. Godager and Wiesen (2013) examines the marginal rate of substitution between earnings and patient health benefits for 42 medical students. Altruism was found to be important, and the majority of students placed more weight on altruism than profit. Crea et al. (2019) has studied the extent to which physicians' altruism is reflected in prescriptions (generic versus branded), using a panel register containing Finnish prescription data from 2003 to 2010. They found no evidence that the probability of prescribing a generic rather than a branded equivalent is affected by any patient characteristics or the out-of-pocket costs faced by the patient. Other studies have examined heterogeneity between physician's responses to monetary motivation in an indirect way by using proxy variables in surveys (see Rizzo and Zeckhauser (2003); Rizzo and Zeckhauser (2007); Iversen and Lurås (2000), among others). Our paper is built on a similar idea and explores how prosocial motivation interacts with the design of social policies. So far, studies of the unintended consequences

⁵Diagnosis Related Groups (DRG) is a hospital output measure approaching complexity of patients' ill episode.

of social policies have focused mainly on taxation and education, where the benefits are obtained by the agent who is behaving strategically (for instance, Chetty and Saez (2003) Jacob and Levitt (2003), Dee and Rockoff (2016)). Instead, the evidence on prosocial motivation, where the agent behaves strategically to benefit a third-party, is limited. In education, for example Diamond and Persson (2016) have detected a form of pro-social motivation in test-score manipulation as teachers compensate for a “bad test day”. On a similar topic, Lavy and Sand (2015) found that teachers grade boys more favourably, and that this increases subsequent enrollment in advanced courses. These papers, similarly to our research, identify social-policy situations where a disadvantaged group is the beneficiary of a strategic decision made by someone else. This highlights a second contribution we make to the literature, on the unintended consequences of public policies, i.e. that the results of individual’s behaviour could take the outcome away from that predicted using assumptions that do not consider this effect. Furthermore, our paper also contributes to the methodological literature on the counterfactual estimation in the presence of notches, using non-parametric methods where the kink regression method is inappropriate given that the assessment behaviour is not randomly distributed near the cut-offs.

The remainder of this paper is organised as follows. Section 2 describes the Spanish LTC system, presenting the institutional setting and the data. Section 3 presents our hypothesis and the theoretical model. In Section 4, we quantify the unintended financial consequences of this prosocial behaviour. Section 5 explores the implications of the number of thresholds in the setting. Finally, section 6 concludes.

2 Long-Term Care system in Spain

2.1 Institutional Framework

In 2006, the Spanish Government approved the *Act to Promote Personal Autonomy and Care* establishing a universal system of Long-Term Care subsidies.⁶ LTC is defined as the permanent assistance required to perform activities of daily living (ADL) required by persons who suffer from a reduction of functional capability.⁷ Despite the universality of the coverage, the system allocates resources on a needs basis, categorising claimants into six different grades of LTC (see Figure A.1, in the

⁶The Act 39/2006 on December, 14 ((BOE, 2006)).

⁷Activities of Daily Living consist of *Basic* activities –such as eating, dressing, bathing, getting in and out of bed, using the toilet and continence– plus supplementary *Instrumental* activities –which include preparing own meals, cleaning, laundry, taking medication, getting to places beyond walking distance, shopping, managing financial affairs and using the telephone/internet–.

Appendix). Grades are increasing in the intensity of care (i.e. the number of hours of support or the value of subsidy), according to the severity of autonomy lost. There is a broad set of subsidised care offered, from institutions such as nursing homes and day care centres, professional home care services, tele-assistance and subsidies for informal caregivers. Figure 2 summarises the average monthly subsidy provided in each grade; (Table A.1, in the Appendix, provides the exact numbers for the different care options).

The severity of needs is calculated by HCPs using a scale consisting of 47 tasks of basic and instrumental activities of daily living. These tasks are grouped into nine different categories: (i) eating and drinking, (ii) control of physical needs, (iii) bathing and hygiene, (iv) other physical care, (v) mobility, moving inside the home, (vi) moving outside the home, (vii) maintaining one’s health and (viii) housework; additionally for those with cognitive impairments, (ix) the capacity to take decisions is also assessed.⁸ For each task, the assessment is made of: (i) the type and frequency of assistance required (i.e. the type ranges from supervision to replacement, and frequency from almost never to always) and (ii) the source of the impairment, namely whether it is a physical limitation, a mental limitation (understanding and coherent performance) or both. The outcome of this needs’ assessment produces an additive score, which is mapped into a grade as follows:

- from 0 to 24 not eligible for public LTC benefits
- from 25 to 39, Grade I-I, giving access to the first menu of benefits
- from 40 to 49, Grade I-II, giving access to the second menu of benefits
- from 50 to 64, Grade II-I, giving access to the third menu of benefits
- from 65 to 74, Grade II-II, giving access to the fourth menu of benefits
- from 75 to 89, Grade III-I, giving access to the fifth menu of benefits
- from 90 to 100, Grade III-II, giving access to the sixth menu of benefits.

2.2 Data

The administrative database used in this study is drawn from the Catalan Institute of Care and Social Services (ICASS). It consists of records of all individuals who claimed LTC benefits in Catalonia, the Northeast Spanish region, between 2008 and 2011. It contains 361,292 individuals and

⁸The scale is called Barem de Valoració de la Dependència, BVD, and is defined in the Royal Decree 504/2007 ((BOE, 2007))

includes information on their socio-demographic characteristics (age, gender, marital status, zip code, health-care records, disability status, cognitive impairments, etc.) and their application form (application dates, team performing the needs assessment, LTC score, LTC grade, and benefits in cases where the claimant was deemed eligible). Table 1 reports the descriptive statistics of claimant characteristics by LTC grade.

The sample is dominated by women (66%), widows (41%) and individuals resident in Barcelona (73%). On average, applicants are 79 years old. Around 25% of the sample receive disability benefits and almost half of the sample (46%) have cognitive impairments. Their average annual income is around 11,000 euros. 31% of claimants were classified in Grade III, 26% in Grade II, 21% in Grade I and the remaining 22% were found to be ineligible for benefits.

The distribution of LTC scores assigned to the claimants is reported in Figure 1.⁹ We observe that the distribution contains kinks around the six thresholds that determine the level of LTC benefits. In the next section, we discuss why we believe that prosocial motivation of HCPs causes these kinks.

3 Hypothesis and Model

3.1 Prosocial Motivation Evidence

Why do HCPs adjust claimant scores around the threshold as reported in the Figure 1? There are several aspects of the assessment process that rule out the possibility that HCPs' behaviour is driven by economic incentives. HCPs are not directly employed by the social services authorities that provide and manage the LTC system. They are organised at local level and receive a fixed wage, with no financial incentive for the quality/quantity of assessments they perform. There is a protocol that they are expected to follow but no auditing of their assessments. Unlike general practitioners (GPs), HCPs do not have a repeated relationship with claimants, meeting just for their one-off assessments. With claimants only informed of the grade (not the score) at a later date by the social services authorities, there should be no pressure on HCPs to adjust the score during the process. Moreover, HCPs have to medically justify their decisions regarding claimants' mental and physical limitations across the 47 tasks that are assessed to obtain the score. In particular, they have to list

⁹The Figure 1 represents the distribution of scores in 2011. The distributions of the years 2008-2010 have a similar pattern and these Figures are available upon request.

the medical diagnosis registered in the Spanish NHS records that leads to each limitation identified. Whether the characteristics of HCPs could affect the assessment is not possible to test, because we do not have any information about them. However, the assessment process is designed in a way that restricts the opportunity of HCPs to manipulate the scores. Each assessment is allocated to a randomly assigned examiner, and the results of the examination are discussed and agreed within a team of HCPs before the score is sent to social services and then the patient is informed of the grade of support they will receive – but not the actual score awarded in their assessment.

The fact that assessments are agreed by teams will limit the influence of any differences at individual level and the distributions of scores across the assessment teams (23 in total in Catalonia) supports this view (see Figure 3). The distributions all have similar shapes despite demographic and health differences across the local authority areas (Observatori del Sistema de Salut de Catalunya, , 2015). Furthermore, the same kinks are also identified in another Spanish region. Figure 4 presents the distribution of scores in the Basque Country, and the peaks around the care-level thresholds are evident.¹⁰ This evidence suggests that even if there are differences between HCPs in their assessment and adjusting behaviour, the approach proposed in the next section is valid at a team level.

Moreover, using the Spanish Disability and Dependency Survey for 2008 (SDDS) we can show that the discontinuity is not representative of the real LTC needs distribution of the Spanish population. The SDDS includes questions on limitations that are very close to the 47 tasks in the basic and instrumental ADL assessment used by the HCPs. These questions were used by Vilaplana Prieto to compute the LTC need score for the individuals who answered the survey, and their needs distribution does not feature any notches around the thresholds of different grades (see Figure 5).¹¹ Thus, we claim that the discontinuities are linked to policy implementation.

Prosocial motivation, common amongst public workers (see Besley and Ghatak (2005); Dur et al. (2014)), could explain adjusting behaviour in the needs assessment and therefore be responsible for the shape of the needs distribution. There is solid evidence that HCPs are concerned about their patients' utility and that they are willing to adjust assessed needs (when a threshold determines the level of care) if this could increase the amount of care received (Ellis et al. (1986) and Almond et al. (2010)). However, given the protocols associated with assessment (previously described), there is not room for large adjustments. Yet, there is evidence that lying has a cost (Gneezy (2005); Fischbacher

¹⁰We are unable to perform further analysis on this region because we have data on just the score assigned and the total number of applicants.

¹¹The spike around 0 is because the survey is representative of the Spanish population and thus includes a large number of people with full-functional capability and autonomy.

and Heusi;Erat and Gneezy (2009)), and as such prosocial HCPs do not exaggerate patient needs universally in a manner that would shift the distribution of assessment scores to the right. They only adjust those individuals at the margin, who could benefit from the modest/limited adjustment that the system allows.

We also observe that examiners are responsive to policy changes. In 2012, a reform of the LTC Act reduced the number of care grades from six to three (BOE, 2012). The distribution of scores after the reform does not demonstrate any clustering around the suppressed thresholds. By 2014, the distribution of scores had become smooth around the suppressed cut-offs, and at the same time, the notches at remaining cut-offs increased substantially as depicted in Figure 6).¹² We also rule out the possibility that re-assessment (increasing future workload) could incentivise the adjustments because the number of reassessments is very low, and usually takes place to account for deterioration in the health of individuals: only 27% of claimants between 2008 and 2011 applied for a reassessment with most concentrated in grade III, in cases where the health status of the claimant is deteriorating rapidly.

To explore whether the adjusting behaviour could be explained by heterogeneity between claimants who are scored close to each threshold, we run a linear probability model for claimants 3 points above and 3 points below each threshold. Results are reported in Table 2, and suggest that their observable characteristics do not explain the adjustment. The coefficients estimated for the observable characteristics –such as age, gender, disability, marital status and chronic condition – are not significantly different for individuals just above and just below the thresholds for the different grades of benefits, after controlling for province, year and assessment-team fixed effects. We have also explored whether there are features of specific types of patient that can explain the kinks by using the same regression model on subsamples of our dataset. These groups are created using factors that are proven to affect LTC and which therefore could potentially affects examiners' behaviour. In particular, medical conditions which are strongly linked to LTC needs, receipt of disability benefits and marital status (see de Meijer et al. (2011), for instance). The estimated differences for claimants above and below the thresholds remain insignificant and small among these groups.¹³ A possible explanation why we do not find heterogeneity among claimants around the kinks could come from the way in which the scale for assessing LTC severity is designed. As explained above, the assessment performed by HCPs is based on 47 tasks, without accounting for personal characteristics of the applicants. The scale does not account for type of residence, cohabitation, financial capacity or

¹²Figure 4 shows the same pattern in the Baske Country.

¹³Results are available upon request.

several other factors that have some effect on the loss of autonomy.

Finally, we rule out the possibility that reputational concerns of HCPs could lead to the manipulation of assessment results because of two significant features of the assessment process, namely: (i) there is only one encounter between the HCP and the claimant, and (ii) the score obtained in the assessment is not directly communicated to the claimant (i.e. is not public information).

3.2 Theoretical model

We now present a theoretical framework in which we can identify the unintended consequences caused by the behavior of HCPs in a non-linear benefits system when there are no financial incentives for HCPs.

We consider a standard model, where the individual's utility depends on Health and Income $U(Y, H)$. For an individual with LTC needs, d , her health status and income are linked to a score that identifies the severity of her needs (θ). Thus, the utility function can be expressed as $U_d(b(\theta), H(\theta))$, where $H'(\theta) < 0$ and $b'(\theta) > 0$, and rewritten as:

$$U_d(b_d, \theta_d) = \gamma b(\theta_d) - (1 - \gamma)\theta_d \quad (1)$$

where γ ($0 < \gamma < 1$) is the weight for each component health and income, (θ and b). We assume that the severity of LTC needs is uniformly distributed among the population; $\theta \sim U(0, 1)$, and $f(\theta)$ is the associated density function.¹⁴

Policy-makers allocate benefits to LTC claimants based on the extent of their needs. As needs are not directly observable to policy-makers, they are measured by HCPs in an assessment process or test. The HCP is assumed to be interested in both her profits and the benefits awarded to her patients/claimants (Ellis and McGuire, 1986). The HCP receives a fixed salary so the components of her utility function can be simplified to the claimant's utility, and the costs incurred for adjusting the assessment directly depends on the number of points added to the score (x_{hcp}), and can be

¹⁴We could assume that this function is a transformation from a conventional Cobb-Douglas function ($U(Y, H(\theta)) = Y^\gamma H^{(1-\gamma)}$). By applying a log-transformation, we could find the utility presented: $U(Y, H(\theta)) = \gamma \log(Y) + (1 - \gamma) \log H(\theta) = \gamma b(\theta) - (1 - \gamma)\theta$, where $\log Y = b(\theta)$ and $\log H(\theta) = -\theta$. Note that health (H) increases claimants' utility but the level of LTC needs θ decreases her utility. We assume that the benefits b increase with the disability score (θ), as the amount of benefits grows with needs.

expressed as:

$$U_{hcp}(U^d(b_d(\theta_d + x_{hcp}), \theta_d), C(x_{hcp})) = \alpha(\gamma b_d - (1 - \gamma)\theta_d) - \left(\frac{x_{hcp}^2}{2}\right) \quad (2)$$

where α ($\in [0, 1]$) accounts for the prosocial preferences of the HCP. When $\alpha > 0$ the *HCP* could marginally adjust the score, θ_d , increasing it by x_{hcp} points: $\theta_d^{hcp} = \theta_d + x_{hcp}$. The adjustment cannot exceed a certain amount, ($0 \leq x_{hcp} \leq \epsilon$, where $0 < \epsilon < 1$).¹⁵ The adjustment imposes a convex cost, generating disutility for the HCP: $C(x_{hcp}) = \frac{x_{hcp}^2}{2}$. There is a risk that an adjustment may be discovered by the wider authorities and if so, the assessor responsible is likely to be reprimanded and suffer a personal cost. This means that HCPs will have an aversion to lying. This cost is likely to grow as the adjustment increases, because a larger adjustment to the score will increase the chance of the lie being discovered –so that the cost function (θ_d^{hcp}) is convex.

Assuming that the policymaker has fixed a threshold of LTC needs, claimants with needs below the threshold, denoted by θ_t , receive y , while claimants with needs above it, receive $y + \Psi$, as specified below:

$$b_d = \begin{cases} y & \text{if } \theta_d^{hcp} \in [0, \theta_t) \\ y + \Psi & \text{if } \theta_d^{hcp} \in [\theta_t, 1] \end{cases}$$

As the utility of the HCP depends on their prosocial preferences and the benefits that claimants could receive, any decision to adjust claimant assessments will be affected by the structure of the LTC benefits available. We distinguish two scenarios:

(a) The *HCP* does not have prosocial preferences. If examiners do not have *prosocial preferences* ($\alpha = 0$), no adjustments take place: $x_{hcp} = 0 \forall \theta_d$, and the adjustment cost is zero. This situation is associated to a budget of LTC:

$$Budget_0 = y \left(\int_0^{\theta_t} f(\theta_d) d\theta_d \right) + (\Psi + y) \left(\int_{\theta_t}^1 f(\theta_d) d\theta \right) = y + \Psi(1 - \theta_t) \quad (3)$$

(b) The *HCP* has prosocial preferences. If examiners have *prosocial preferences* ($0 < \alpha \leq 1$), score adjustment could happen in the assessment process. As previously discussed, adjusting the

¹⁵The adjustment can only be a marginal adjustment for two reasons. First the medical conditions of the claimants need to be certificated, and secondly, even if there is some personal cost, it is well documented that people tend to tell the truth (Gneezy (2005); Fischbacher and Heusi; Erat and Gneezy (2009)).

score is costly to the HCP, and this cost is higher when the manipulation is bigger, i.e. when the applicant's score is far from the threshold that defines the benefits. The *HCP* will take one of three decisions driven in part by the (*true*) score of the applicant (θ_d): HCP compares her U_{hcp} if she adjust (*a* denotes adjusting response: i.e. U_{hcp}^a) and if she does not adjust (*na* to denote not-adjusting response: i.e. U_{hcp}^{na}). The HCP assesses the decisions of adjusting for each type of claimant need (θ_d):

- **(i)** $\theta_d \geq \theta_t$: The adjustment does not increase the patient's benefits and the *HCP* incurs only a positive cost, thus the adjustment is not made. Formally:

$$U_{hcp}^a \leq U_{hcp}^{na} \iff \alpha(\gamma(y + \Psi) - (1 - \gamma)\theta_d) - \frac{(x_{hcp})^2}{2} \leq \alpha(\gamma(y + \Psi) - (1 - \gamma)\theta_d)$$

- **(ii)** $\theta_d < \theta_t$ and the patient's score (θ_d) is far from the threshold (θ_t): $\theta_d^{hcp} = \theta_d + x_{hcp} < \theta_t$. Thus, the cost of adjusting is larger than the benefit from adjusting, so that the adjustment is not made. Formally:

$$U_{hcp}^a \leq U_{hcp}^{na} \iff (\alpha(\gamma y - (1 - \gamma)\theta_{dhcp}) > \frac{(x_{hcp})^2}{2}) \leq \alpha(\gamma y - (1 - \gamma)\theta_d)$$

- **(iii)** $\theta_d < \theta_t$ and the patient's score (θ_d) is close to the threshold (θ_t): $\theta_d^{hcp} = \theta_d + x_{hcp} \geq \theta_t$ (where $0 < x_{hcp} \leq \epsilon$), the adjustment will take place. Formally:

$$U_{hcp}^a > U_{hcp}^{na} \iff \alpha(\gamma(y + \Psi) - (1 - \gamma)\theta_{dhcp}) - \frac{(x_{hcp})^2}{2} > \alpha(\gamma y - (1 - \gamma)\theta_d)$$

$$\text{which occurs if } \frac{(x_{hcp})^2}{2} = \frac{(\theta_{dhcp} - \theta_d)^2}{2} \leq \alpha(\gamma y - (1 - \gamma)\theta_d)$$

A graphical representation of these three cases is presented in the Appendix (see Figure A.2). The number of individuals whose scores are adjusted depends positively on: the prosocial preferences of the HCP ($\frac{\partial x_{hcx}}{\partial \alpha} > 0$), the difference in the value of benefits between categories ($\frac{\partial x_{hcx}}{\partial \Psi} > 0$), and the weight assigned to the LTC benefits in the HCP utility function ($\frac{\partial x}{\partial \gamma} > 0$). Assuming that all examiners have the same prosocial preference level, their decision is equivalent to a shift of the threshold towards the left, denoted by $\theta_{t'} = \theta_t - \epsilon$. Applicants' whose score is $\theta_d \leq \theta_t - \epsilon$ would all be adjusted as illustrated in Figure A.3 in the Appendix. The larger the ϵ , the more applicant scores will be adjusted. In this scenario the associated budget of LTC is:

$$Budget_1 = y \left(\int_0^{\theta_{t'}} f(\theta_d) d\theta_d \right) + (\Psi + y) \left(\int_{\theta_{t'}}^1 f(\theta_d) d\theta_d \right) = y + \Psi(1 - \theta_{t'}) \quad (4)$$

Compared to the non-prosocial preferences case, the budget required is greater $\Delta Budget = \Psi(\theta_t - \theta_{t'})$.

4 Financial Consequences of Prosocial Behavioural

Economists such as Smith (1759), Becker (1974) or Samuelson (1993), note that people often care for the well-being of others and this tendency may have important economic consequences. In this Section, we estimate how many claimants have been adjusted to help quantify the financial public cost that occurs as a result of adjusting scores in the LTC assessment.

To quantify the monetary cost of these adjustments around the thresholds we first need to estimate how many people's scores have been adjusted. To do so we have implemented a non-parametric method to reconstruct the true distribution of the scores. In particular, we use a geometric progression (also known as a geometric sequence analysis). Then, by undertaking "back of the envelope" calculations we compute the costs associated with the pro-social behavior of HCP.

Looking at the distribution of the scores around the thresholds, we can see that a geometric progression is an appropriate way to model the linear function in these areas of the distribution. A geometric progression consists of a sequence of numbers where each term (after the first) is found by multiplying the previous one by a fixed, non-zero number called the common ratio (λ) (see Hazewinkel (2001) for more details). This method is widely used in statistics to forecast population growth (Stanbery, 1952). The basic model for geometric change in population size is: $P = P_0\lambda^t$. We employ a similar technique to recover the true distribution of the population around the thresholds. To recover the frequency of claimants below the threshold n_0 (for example 24) we need to reallocate some of the observations at the scores above n_0 (for example 25, 26, etc.) to the n_0 score. Defining the true frequencies (i.e. the number of applicants in each score) as: $n_0, n_1, n_2, n_3 \dots n_k$, and applying the geometric progression rule we can write the distribution of the scores as:

$$\begin{aligned}n_1 &= (1 - \lambda)\theta_1 \\n_2 &= (1 - \lambda^2)\theta_2 \\n_k &= (1 - \lambda^k)\theta_k \\n_0 &= \theta_0(1 + \lambda + \dots + \lambda^k)\end{aligned}$$

Assuming that the parameter (θ) is the same for each score n we can rewrite the equations

above for each score as:

$$\begin{aligned}
n_{s+1} &= (1 - \lambda^{s+1})\theta \\
n_{s+2} &= (1 - \lambda^{s+2})\theta \\
n_{s+k} &= (1 - \lambda^{s+k})\theta \\
n_s &= \theta(1 + \lambda^s + \dots + \lambda^{s+k}) \\
&= \theta \frac{1 - \lambda^{s+k+1}}{1 - \lambda}
\end{aligned}$$

Looking at this distribution of the score n_s , it approximates to a Poisson distribution: $f(\mu, k) = e^{-\mu} * \mu^{n_k} / k!$. Where $\mu = \theta(1 - \lambda^i)$. Making the assumption that k is five, i.e. two points above and two points below the threshold, the associated likelihood to this function could be expressed as:

$$L(\theta, \lambda) = \frac{1}{n_0!} \exp \left\{ -\theta \left(\frac{1-\lambda^5}{1-\lambda} \right) \right\} \left\{ \theta \left(\frac{1-\lambda^5}{1-\lambda} \right) \right\}^{n_0} \prod_{i=1}^4 \frac{1}{n_i!} \exp \left\{ -\theta (1 - \lambda^i) \right\} \left\{ \theta (1 - \lambda^i) \right\}^{n_i}$$

and its log-Likelihood is

$$\begin{aligned}
l(\theta, \lambda) = \log L(\theta, \lambda) &= -\theta \left(\frac{1 - \lambda^5}{1 - \lambda} \right) + n_0 \log(\theta) + n_0 \log \left(\frac{1 - \lambda^5}{1 - \lambda} \right) + \\
&\quad \sum_{i=1}^4 -\theta (1 - \lambda^i) + n_i \log(\theta) + n_i \log (1 - \lambda^i)
\end{aligned}$$

$$\frac{\partial l(\theta, \lambda)}{\partial \theta} = - \left(\frac{1 - \lambda^5}{1 - \lambda} \right) + \frac{n_0}{\theta} + \sum_{i=1}^4 \left(- (1 - \lambda^i) + \frac{n_i}{\theta} \right)$$

$$\frac{\partial l(\theta, \lambda)}{\partial \lambda} = - \left(\frac{5\theta\lambda^4 - \theta(1-\lambda^5)(1-\lambda^{-1})}{1-\lambda} \right) + \frac{n_0}{1-\lambda^5} \left(-5\lambda^4 + \frac{1-\lambda^5}{1-\lambda} \right) + \sum_{i=1}^4 \left((i\theta\lambda^{i-1}) - \frac{i\lambda^{i-1}n_i}{1-\lambda^i} \right)$$

To maximize this function we use a Newton-Rapson estimator and can calculate θ and λ close to each threshold. After identifying θ and λ and assuming homogeneity among individuals around the threshold, we calculate the number of people who have been adjusted in each score (see A.2, by years).¹⁶ We estimate that almost 9000 claimants (3%) between 2008 and 2011 had their score adjusted. Using these values, we are able to create a counterfactual distribution of the scores. This is reported in Figure 7.¹⁷ To ensure the validity of our approach, we test whether this methodology produces values similar to the populations at scores far from the thresholds in the observed distribution (see Figure 8). This figure shows that the non-parametric model developed produces values close to the original numbers of claimants at all scores not affected by the adjustments.

To estimate the impact on LTC public spending, we perform “back of the envelope” calcula-

¹⁶As Table 2 shows, we do not find a notable difference between claimants’ scored below and above each threshold.

¹⁷We present here only the distribution of 2010, the other years in the study are available upon request.

tions to compute the cost of adjusting across all those claimants who were receiving voucher benefits in November 2010.¹⁸ From all claimants scored 3 points or less above each threshold of each year, we selected at random the individuals who we label as “adjusted” to create a counterfactual distribution. From the adjusted individuals we identify those who received a voucher and then calculate the aggregate value of these vouchers to estimate that approximately 838,136 euro per month extra is spent on benefits as a result of 9000 claimants being adjusted by HCPs’ pro-social behaviour. (see Table 3). This cost is obtained by multiplying the number of adjusted claimants who receive a voucher or cash transfer benefit in each category by the average difference in expenditure per claimant above and below the threshold. This cost is underestimated, because we use the value of the voucher awarded, not the costs incurred by the service provider as this information is not provided. In addition, we perform the calculation without knowing the family income of the applicant, a factor which is taken into account when deciding the benefits awarded, nor any in-kind benefits which we cannot place a financial value on. This back of the envelope calculation also assumes that these adjusted claimants would choose the same type of benefit if they were placed in the lower category of support.

5 Does the number of thresholds matter?

Debate in public policy has suggested that smooth-functions should be adopted to correct the type of unintended consequences of step-functions that we identify here when allocating public benefits (see Blomquist et al. (2015), for non-linear budgets and when considering taxable income; Van Doorn et al. (2007) on attitudes and behaviour of decision makers, and Einav et al. (2017) for spending responses to health insurance contracts. Motivated by these findings, we discuss how the number of grades (and therefore, the number of thresholds) could affect the adjusting behaviour. On one side it might be that an increase of the number of thresholds, which reduces the differences in benefits across grades, could reduce the incentive to adjust. However, the decision to adjust is only relevant for those claimants close to the thresholds, so if the number of thresholds reduces the number of people “eligible” for adjustment decreases (as we can deduct from Section 3). Given these counteracting issues, we cannot conclude from the theory whether more thresholds could help to reduce the unintended consequences of the adjusting behavior.

¹⁸From all of the claimants assessed entitled with benefit only 46% were actually receiving. From those, almost 70% received a financial benefit. Claimants who receive an in-kind benefit are excluded from this analysis because obtaining the cost of the service they benefit is not possible.

If we add more grades into our benefits function presented in section 3.2., we would identify that not only people with LTC needs just below θ_t would be adjusted but also people just below the new threshold, call it θ_{t^*} . Yet the amount of people adjusted and the budget implications will depend on the parameters that define our model.

To shed light on this debate, we explore empirically a 2012 policy change that saw the Spanish Government reduce the number of grades (and therefore the number of thresholds), intending to reduce public spending in line with the fiscal austerity caused by the Great Recession.

Following the same methods applied in the Section 4, we estimate the counterfactual score distribution of 2014 to calculate the number of adjusted claimants in both scenarios. Table 4 provides the number of adjusted scores for individuals in 2014 compared to the 2010 adjustments, with three thresholds in 2014 and six in 2011. We find that the total number of adjustments is higher with more thresholds (5.6% of claimants versus 5.3% of claimants, in 2011 and 2014 respectively), although the difference in claimants' adjustments is small (only 0.3 pp). This empirical finding suggests that the two opposite forces compensate, and the adjustment remains constant in both scenarios. However, from our theoretical model we learn that the probability of being adjusted depends on the proximity to the threshold, i.e. it is important to consider claimants "eligible" for adjustment. If we focus on the percentage of adjustment around the thresholds that were retained after the reform (i.e. 25, 50 and 75), the number of adjustments was bigger after the reform (i.e. the percentage of adjustment is 3.6% and 5.3%, in 2011 and 2014 respectively). Furthermore, if we only consider the most marginal patients, defined as those scored at or within two points of the six original threshold scores, we find that in 2011 15.9% were adjusted but this rose to 26.9% in 2014. This indicates that adjustments become larger, picking up claimants further from the threshold than was the case when there were more grades of support available. As we can see in Figure 6 the deviation from the smooth distribution starts further from the thresholds in the reformed system with fewer categories. Our calculation should be taken with caution as we do not account for any other changes that may have occurred in the claimant characteristics or public authorities as a result of other financial austerity measures.

6 Conclusions

This paper provides one of the few empirical pieces of evidence on prosocial motivated health care professionals (HCPs). First, we demonstrate that prosocial motivation affects the behaviour of HCPs.

Second, we identify the unintended consequences that occur when prosocial HCPs use a continuous scale to determine the LTC needs of claimants in a system where there is a discontinuous scheme of public LTC benefits. In this setting, HCPs adjust the score of claimants whose real scores are just below the thresholds to move them up to the next grade, granting access to more valuable benefits. We find that around one million euros of the total annual expenditure is due to adjusting, and that adjustment affects 3% of claimants. Thirdly, we compare the adjusting behaviour in two situations where there is variation in the number of thresholds. We find that the share of adjusted claimants around the thresholds is larger when the number of thresholds is reduced, which places a higher risk of discovery upon the HCP performing the adjustment. Whether the adjustments improve or not the quality of life of these claimants is beyond the scope of this paper, but should be analysed in future research. We suggest that this type of behaviour could also be prevalent amongst other public workers like teachers, judges, etc where pro-social behaviour in agents can have unintended consequences for the public purse. Planners should take into consideration the evidence presented here when considering the use of economic incentives for these types of workers because our results show responses are not always straightforward. This could be particularly important when planners cannot observe the principals and agents in decision making and rely on intermediary agents to determine outcomes. Our evidence shows that these intermediary agents are unlikely to behave in ways that completely eliminate moral hazard problems and the introduction of more complex solutions including increased supervision or financial incentives may be needed to reach the intended outcomes.

References

- Almond, D., Doyle, J. J., Kowalski, A. E., Williams, H., and Williams, H. (2010). Estimating Marginal Returns To Medical Care: Evidence From At-Risk Newborns. *The quarterly journal of economics*, 125(2):591–634.
- Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, 100(401):464.
- Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy*, 82:1063–1093.
- Besley, T. and Ghatak, M. (2005). Competition and Incentives with Motivated Agents. *American Economic Review*, 95(3):616–636.
- Blomquist, S., Kumar, A., Liang, C.-Y., and Newey, W. K. (2015). Individual heterogeneity, non-linear budget sets, and taxable income.
- BOE (2006). Act 39/2006, of 14th December, on Promotion of Personal Autonomy and Assistance for Persons in a Situation of Dependency.
- BOE (2007). Royal Decree 504/2007, of 20th April, which approves the scale to value the dependency situation established by Law39/2006 of 14th December.
- BOE (2012). Royal Decree 20/2012, 13th July, of measures to ensure budgetary stability and promoting competitiveness.
- Charness, G. and Haruvy, E. (1999). Altruism, Equity, and Reciprocity in a Gift Exchange Experiment: An Encompassing Approach.
- Chetty, R. and Saez, E. (2003). Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut. *The Quarterly Journal of Economics*.
- Crea, G., Galizzi, M. M., Linnosmaa, I., and Miraldo, M. (2019). Physician altruism and moral hazard:(no) evidence from finnish national prescriptions data. *Journal of health economics*, 65:153–169.
- de Meijer, C., Koopmanschap, M., dâ Uva, T. B., and van Doorslaer, E. (2011). Determinants of long-term care spending: Age, time to death or disability? *Journal of Health Economics*, 30(2):425 – 438.
- Dee, Thomas S, W. D. B. A. J. and Rockoff, J. (2016). The causes and consequences of test score

- manipulation: Evidence from the new york regents examinations. *National Bureau of Economic Research*.
- Diamond, R. and Persson, P. (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. *National Bureau of Economic Research*.
- Dur, R., Zoutenbier, R., Dur, R., and Zoutenbier, R. (2014). Intrinsic Motivations of Public Sector Employees: Evidence for Germany. *German Economic Review*.
- Einav, L., F. A. and Mahoney, N. (2018). Provider incentives and healthcare costs: Evidence from long-term care hospitals. *Econometrica*.
- Einav, L., Finkelstein, A., and Schrimpf, P. (2017). Bunching at the kink: implications for spending responses to health insurance contracts. *Journal of Public Economics*, 146:27–40.
- Ellis, R. P., McGuire, T. G., Ellis, R., and McGuire, T. G. (1986). Journal of health economics. *Journal of Health Economics*, 5(2):129–151.
- Erat, S. and Gneezy, U. (2009). White Lies. *Rady School of Management*.
- Evans, R. G. (1984). *Strained Mercy: The Economics of Canadian Health Care*. Butterworth-Heinemann.
- Fehr, E. and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*, 11(10):419–427.
- Fehr, E. and Falk, A. (1969). European economic review. *European Economic Review*, 46(4-5):687–724.
- Feldstein, M. S. (1970). Corporate Taxation and Dividend Behaviour. *The Review of Economic Studies*, 37(1):57–72.
- Fischbacher, U. and Heusi, F. Lies in Disguise An experimental study on cheating. *Thurgau Institute of Economics*.
- Francois, P. and Vlassopoulos, M. (2008). Pro-Social Motivation and the Delivery of Social Services. *CESifo Economic Studies*.
- Frey, B. S. and Jegen, R. (2001). Motivation Crowding Theory. *Journal of Economic Surveys*, 15(5):589–611.

- Frey, B. S. and Oberholzer-Gee, F. (1997). The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out. *The American Economic Review*, 87:746–755.
- Gneezy, U. (2005). Deception: The Role of Consequences. *American Economic Review*, 95(1):384–394.
- Godager, G. and Wiesen, D. (2013). Profit or patients’ health benefit? Exploring the heterogeneity in physician altruism. *Journal of health economics*, 32:1105–1116.
- Hazewinkel, M. (2001). The Algebra of Quasi-Symmetric Functions Is Free over the Integers. *Advances in Mathematics*, 164(2):283–300.
- Iversen, T. and Lurås, H. (2000). The effect of capitation on GPs’ referral decisions. *Health economics*, 9(3):199–210.
- Jacob, B. and Levitt, S. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*.
- Jürges, H. and Köberlein, J. (2015). What explains DRG upcoding in neonatology? The roles of financial incentives and infant health. *Journal of Health Economics*, 43:13–26.
- Lavy, V. and Sand, E. (2015). On the origins of gender human capital gaps: Short and long term consequences of teachers’ stereotypical biases. *National Bureau of Economic Research*.
- Le Grand, J. (2003). *Motivation, Agency, and Public Policy*. Oxford University Press.
- Levitt, S. D. and List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *The Journal of Economic Perspectives*, 21:153–174.
- Maxwell, B. G., Wong, J. K., Miller, D. C., and Lobato, R. L. (2014). Temporal Changes in Survival after Cardiac Surgery Are Associated with the Thirty-Day Mortality Benchmark. *Health Services Research*.
- Observatori del Sistema de Salut de Catalunya, (2015). Central de resultats. Àmbit d’atenció primària. dades 2014. barcelona: Agència de qualitat i avaluació sanitàries de catalunya. departament de salut. generalitat de catalunya.
- Perry, J. L., Hondeghem, A., and Wise, L. R. (2010). Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. *Public Administration Review*.
- Rizzo, J. A. and Zeckhauser, R. J. (2003). Reference Incomes, Loss Aversion, and Physician Behavior. *Review of Economics and Statistics*, 85(4):909–922.

- Rizzo, J. A. and Zeckhauser, R. J. (2007). Pushing incomes to reference points: Why do male doctors earn more? *Journal of Economic Behavior & Organization*, 63(3):514–536.
- Samuelson, P. A. (1993). Altruism as a problem involving group versus individual selection in economics and biology. *American Economic Review*.
- Siciliani, L. and Verzulli, R. (2009). Waiting times and socioeconomic status among elderly Europeans: evidence from SHARE. *Health Economics*, 18(11):1295–1306.
- Smith, A. (1759). *The Theory of Moral Sentiments*. London.
- Stanbery, V. B. (1952). *Better population forecasting for areas and communities: A guide book for those who make or use population projections*. US Government Printing Office.
- Van Doorn, J., Verhoef, P. C., and Bijmolt, T. H. (2007). The importance of non-linear relationships between attitude and behaviour in policy research. *Journal of Consumer Policy*, 30(2):75–90.
- Vilaplana Prieto, C. Necesidades no satisfechas de las personas dependientes en espana. *In XVI Encuentro de Economía Pública*.

Tables & Figures

Figure 1: Score's Distribution in 2011

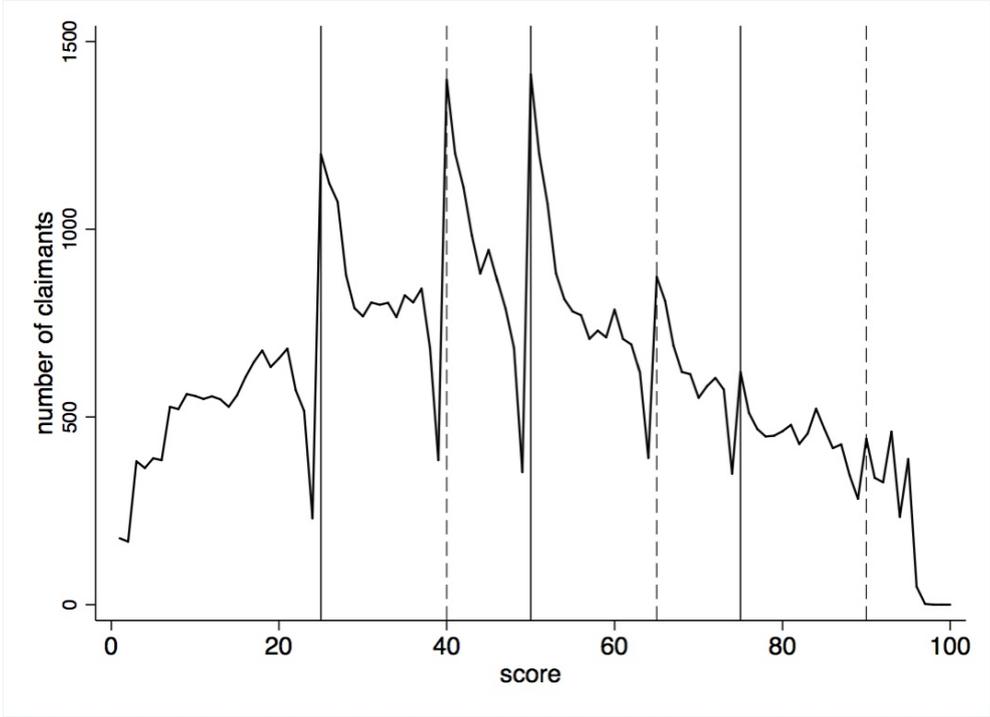


Figure 2: Average Monthly LTC Benefits by Scores

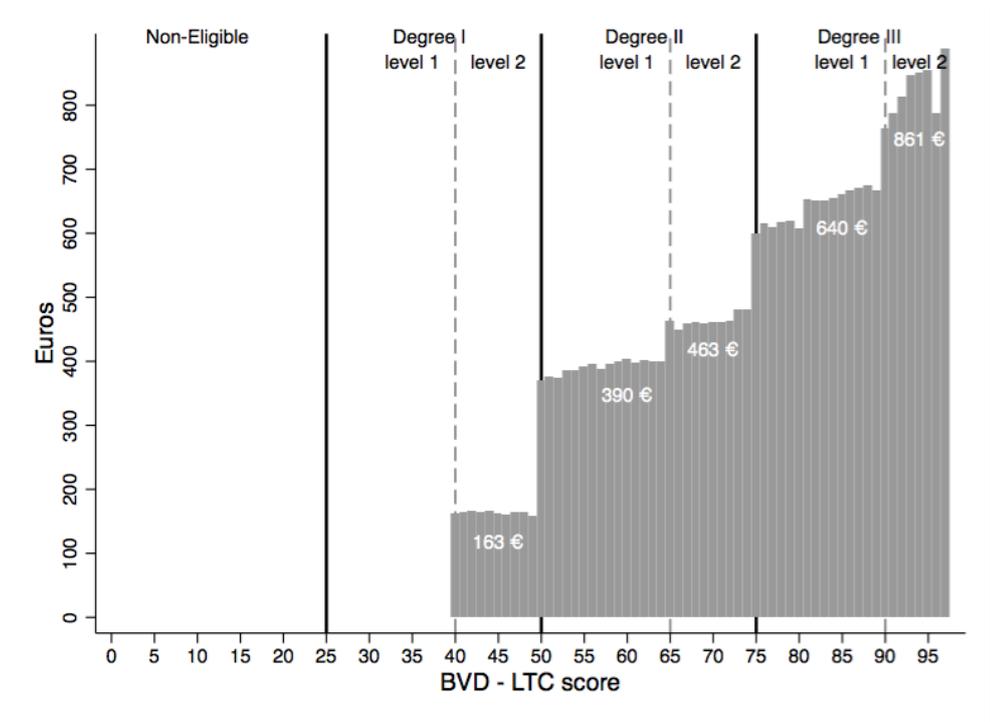


Figure 3: Scores' Distribution by HCP Local Assessment Teams (SEVADs)

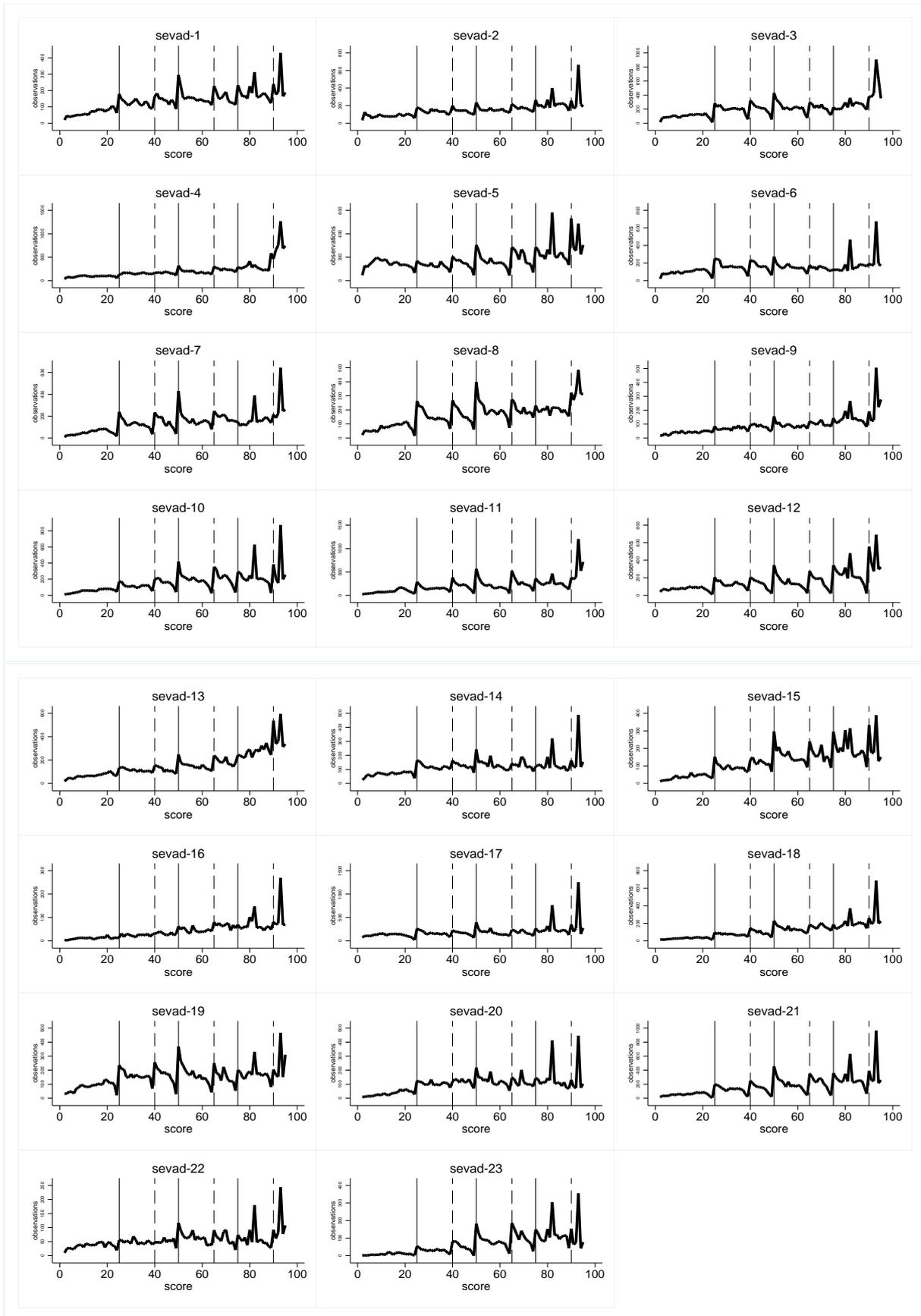
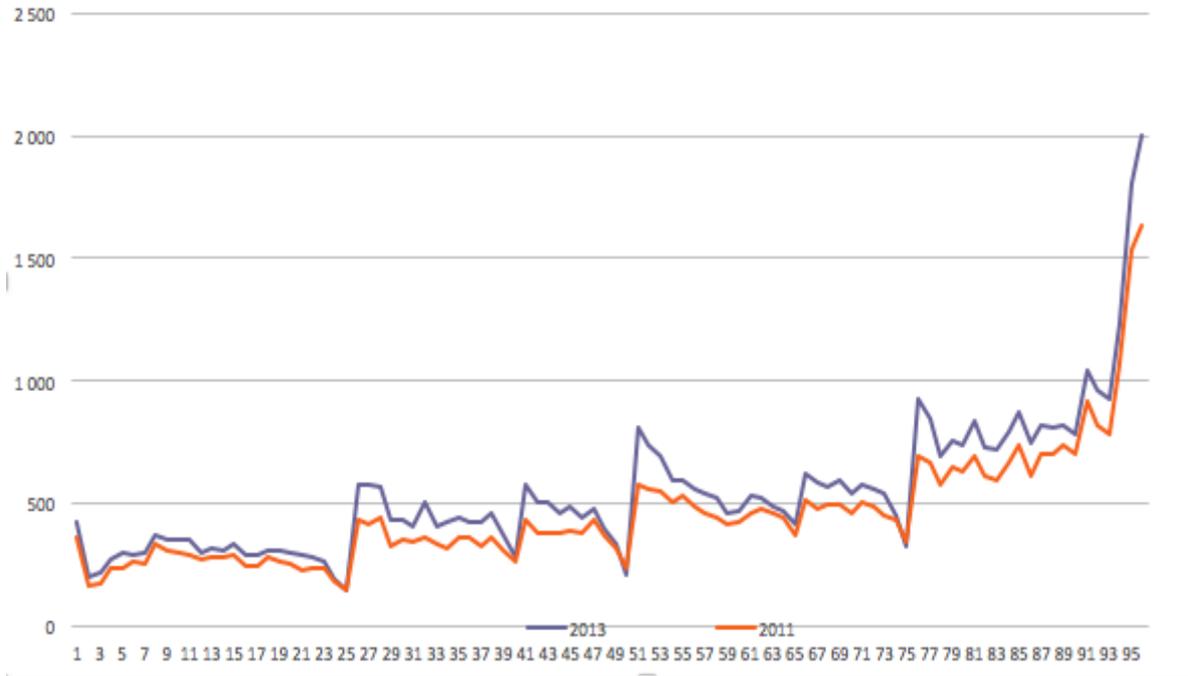
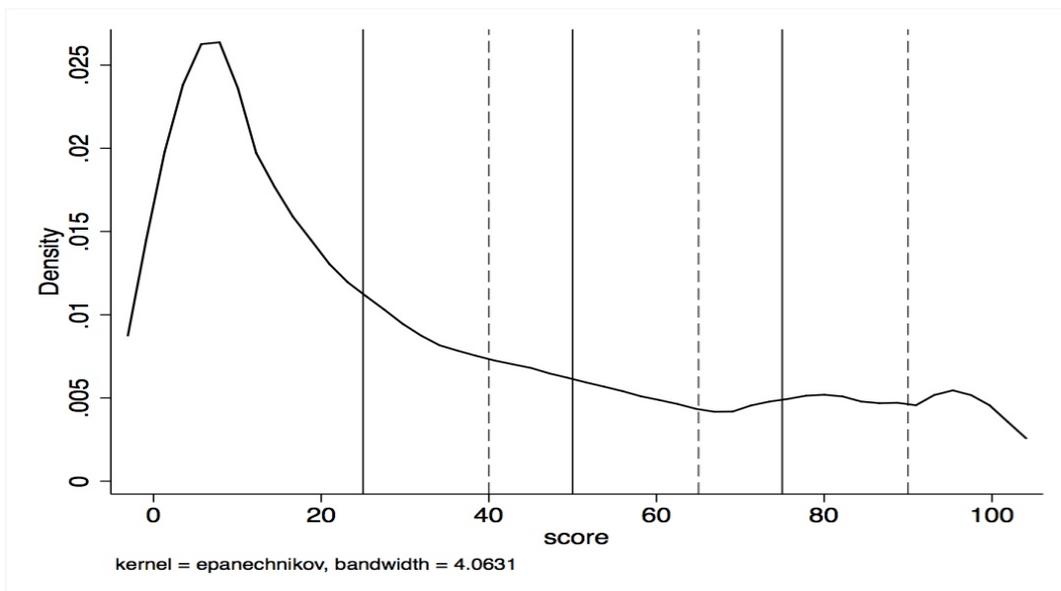


Figure 4: Score Distribution in the Bask Country



Source: Bask Country Region administrative records.

Figure 5: Kernel density of the Estimated Score from SDDS' Survey Responses



Source: Vilaplana, 2010.

Figure 6: Score Distributions Before and After Levels' Removal

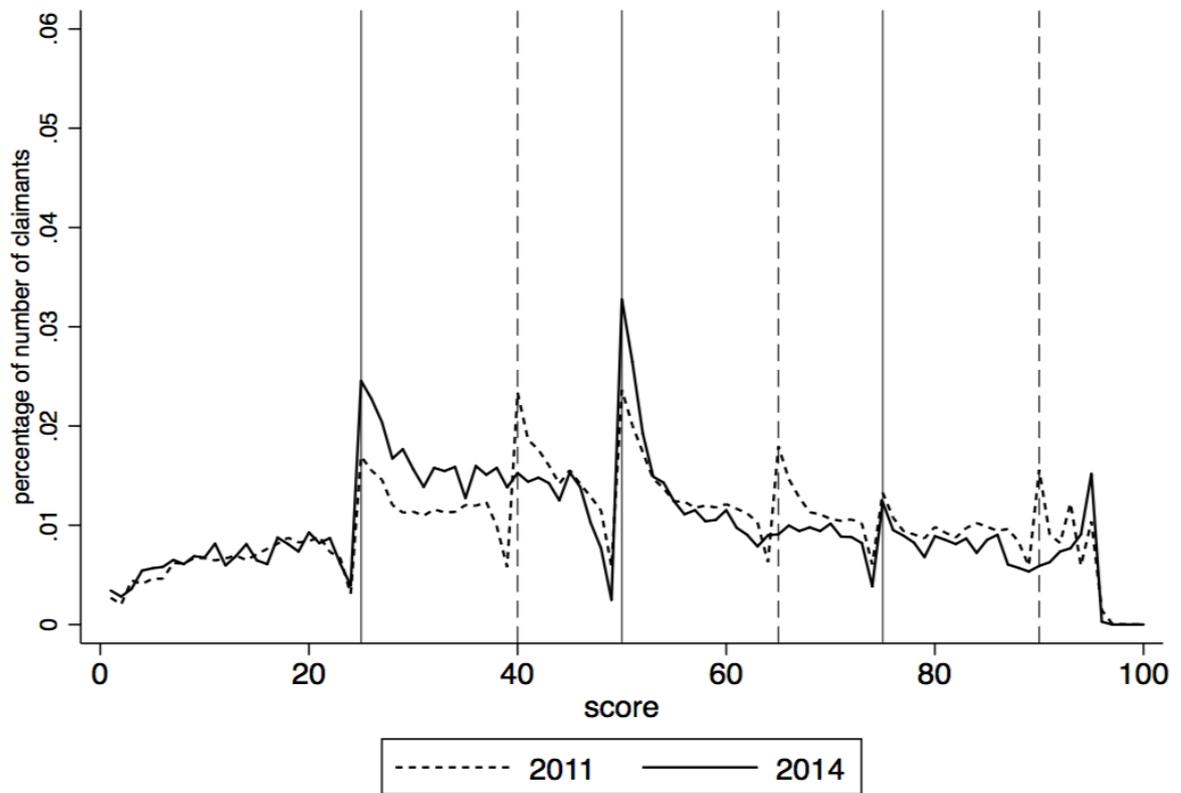


Table 1: **Descriptive Statistics**

	(1)	(2)	(3)	(4)	(5)
	All Sample	No Grade	Grade I	Grade II	Grade III
Female	0.66	0.61	0.69	0.66	0.68
Age	78.99	77.93	77.09	78.37	81.43
<i>Age: 40-54</i>	0.05	0.05	0.07	0.06	0.04
<i>Age: 55-69</i>	0.11	0.13	0.14	0.11	0.07
<i>Age: 70-84</i>	0.48	0.52	0.53	0.48	0.43
<i>Age: +85</i>	0.35	0.30	0.27	0.34	0.45
Civil Status					
<i>Married</i>	0.37	0.42	0.42	0.38	0.30
<i>Widow</i>	0.41	0.38	0.39	0.42	0.45
<i>Single</i>	0.10	0.10	0.11	0.12	0.09
<i>Other CS</i>	0.11	0.09	0.08	0.08	0.16
Region (province)					
<i>Barcelona</i>	0.73	0.75	0.73	0.71	0.73
<i>Girona</i>	0.09	0.07	0.08	0.10	0.10
<i>Lleida</i>	0.07	0.07	0.07	0.07	0.07
<i>Tarragona</i>	0.11	0.10	0.13	0.12	0.11
Year of Application					
<i>2007*</i>	0.15	0.07	0.03	0.08	0.35
<i>2008</i>	0.27	0.20	0.20	0.28	0.33
<i>2009</i>	0.24	0.22	0.29	0.29	0.18
<i>2010</i>	0.20	0.22	0.30	0.23	0.10
<i>2011*</i>	0.14	0.28	0.18	0.11	0.04
Labour Disability	0.26	0.23	0.29	0.28	0.26
Cognitive impairment	0.46	0.24	0.29	0.53	0.76
Annual Income (euros)	11028.62	11629.84	11036.01	10810.79	11295.37
Missing Income	0.61	0.92	0.37	0.41	0.73
Score	52.53	6.39	36.58	61.24	86.13
Access to benefits	0.55	0.00	0.19	0.87	0.83
Voucher or Cash Transfer	0.77		0.69	0.83	0.74
Type of Benefit					
<i>At home Care</i>	0.07		0.13	0.07	0.06
<i>Day Care Centre</i>	0.03		0.03	0.03	0.02
<i>Informal Caregiver</i>	0.63		0.69	0.72	0.54
<i>Medical Nursing Home</i>	0.02		0.01	0.01	0.03
<i>Nursing Home</i>	0.24		0.07	0.15	0.34
<i>TeleAssistance</i>	0.02		0.07	0.02	0.01
Observations	361292	76233	74252	95723	115084
		0.21	0.21	0.26	0.32

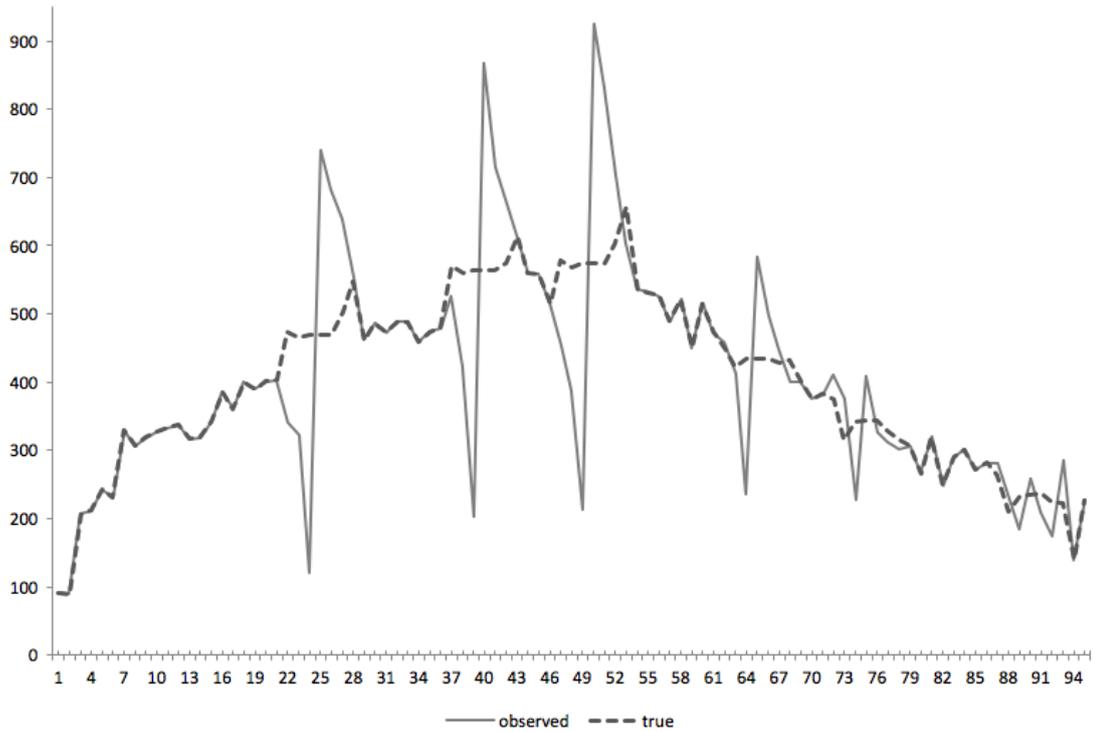
Notes: (*) The implementation in 2007 started in June, and observations of 2011 do not include December.

Table 2: Above the cutoff

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Grade I 11 VS	Grade I 12 VS	Grade II 11 VS	Grade II 12 VS	Grade I 11 VS	Grade I 12 VS	Grade II 11 VS	Grade II 12 VS
	Grade I 12	Grade II 11	Grade II 12	Grade III 11	Grade I 12	Grade II 11	Grade II 12	Grade III 11
	37-42	47-52	62-67	72-77	37-42	47-52	62-67	72-77
Female	-0.014** (0.007)	-0.002 (0.007)	0.003 (0.011)	0.010 (0.011)	-0.016** (0.007)	-0.004 (0.006)	0.004 (0.010)	0.008 (0.011)
Age: 55-69	0.016 (0.020)	-0.004 (0.017)	-0.022 (0.015)	0.013 (0.029)	0.016 (0.020)	0.005 (0.017)	-0.020 (0.015)	0.024 (0.028)
Age: 70-84	0.021 (0.016)	0.022 (0.014)	-0.025* (0.013)	-0.003 (0.022)	0.028 (0.016)	0.036*** (0.012)	-0.012 (0.013)	0.012 (0.020)
Age: +85	0.026 (0.019)	0.026* (0.013)	0.001 (0.016)	0.028 (0.021)	0.038* (0.020)	0.045*** (0.012)	0.018 (0.014)	0.042** (0.018)
Married	0.016 (0.010)	0.001 (0.010)	0.022** (0.008)	0.035** (0.015)	0.006 (0.011)	-0.004 (0.011)	0.018* (0.009)	0.035** (0.015)
Widow	0.030*** (0.010)	-0.001 (0.009)	-0.011 (0.009)	0.004 (0.016)	0.022* (0.012)	-0.006 (0.009)	-0.015 (0.010)	0.008 (0.015)
Disability	-0.009 (0.011)	-0.015 (0.010)	-0.015 (0.009)	-0.014 (0.011)	-0.004 (0.010)	-0.007 (0.009)	-0.003 (0.009)	-0.006 (0.010)
Number of Diagnosis					0.008	-0.007	-0.007	-0.012**
SEVAD F.E.	No	No	No	No	Yes	Yes	YES	Yes
Province F.E	No	No	No	No	Yes	Yes	Yes	Yes
Observations	17,108	23,334	22,317	21,366	17,105	23,324	22,304	21,347

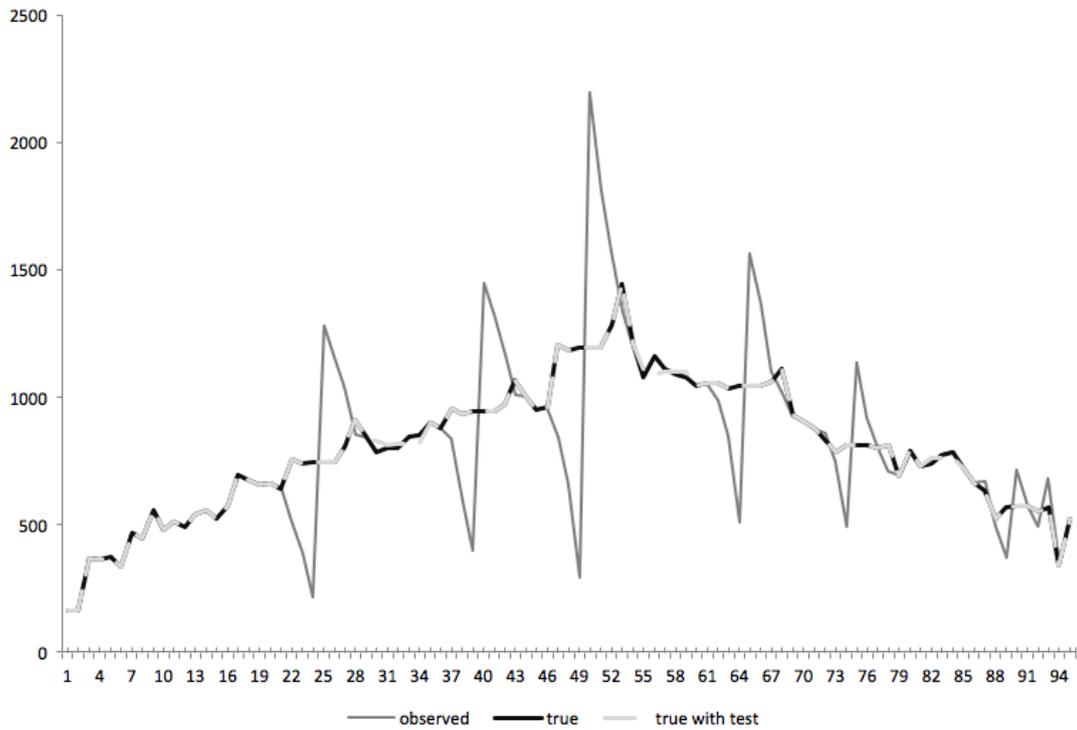
Notes: The dependent variables is a dummy which takes value one if the score is 3 points above the threshold, 0 if 3points below. All OLS regressions include time fixed effects. Robust Standard errors, clustered at SEVAD level. The reference variables are age cohort from 40 to 54 years old and divorced. Significance levels: One star (*) if $p < 0.05$, two stars(**) if $p < 0.01$, and three stars (***) if $p < 0.001$.

Figure 7: Score distributions: observed vs *true*



Notes: This graph presents the distribution of 2010. The *true* distribution is obtain with the counterfactual estimation.

Figure 8: Score distributions: observed vs *true* robustness check



Notes: This graph presents the distribution of 2010. Compared to Graph 7 it includes the *true with test* distribution, which is constructed estimating the counterfactual values in other parts of the distribution which are not affected by the adjustments.

Table 3: **Back of the envelope Calculations**

	(1)	(2)	(3)	(4)	(5)
	Grade I Level II	Grade II Level I	Grade II Level II	Grade III Level I	Grade III Level II
Number of adjusted claimants	31	3424	2395	2063	981
<i>Monthly Expenditure</i>					
With Adjustments	5583	1040948	811797	566136	584345
Without Adjustments	0	605052	723589	471456	442576
<i>Difference or "potential savings"</i>					
Total	5583	435896	88208	94680	141769
Percentage	100.00	41.87	10.87	16.72	24.26

Table 4: **Number of Adjusted Claimants**

		2011	2014
Claimants	All	38235	25249
	Around thresholds	13569	4974
Adjusted claimants	Grade I -level 1	618	511
	Grade I -level 2	540	
	Grade II -level 1	713	699
	Grade II - level 2	229	
	Grade III -level 1	32	127
	Grade III -level 2	25	
	All Adjusted	2157	1337
	Adjusted at 25, 50 & 75	1363	1337
Percentage of adjusted claimants over	total claimants	5.6%	5.3%
	total claimants (only adjusted at 25, 50 & 75)	3.6%	5.3%
	claimants around thresholds	15.9%	26.9%

Notes: "Around the thresholds" includes all claimants scored at the threshold and two scores above and below. Thresholds at 25 (Grade I - level 1), 50 (Grade II - level 1) and 75 (Grade III - level 1) are fixed and unaffected by 2012 reform, which eliminates thresholds at 40, 65 and 90.

Appendix A Supplementary Tables and Figures

Figure A.1: Spanish LTC system: Funnel procedure

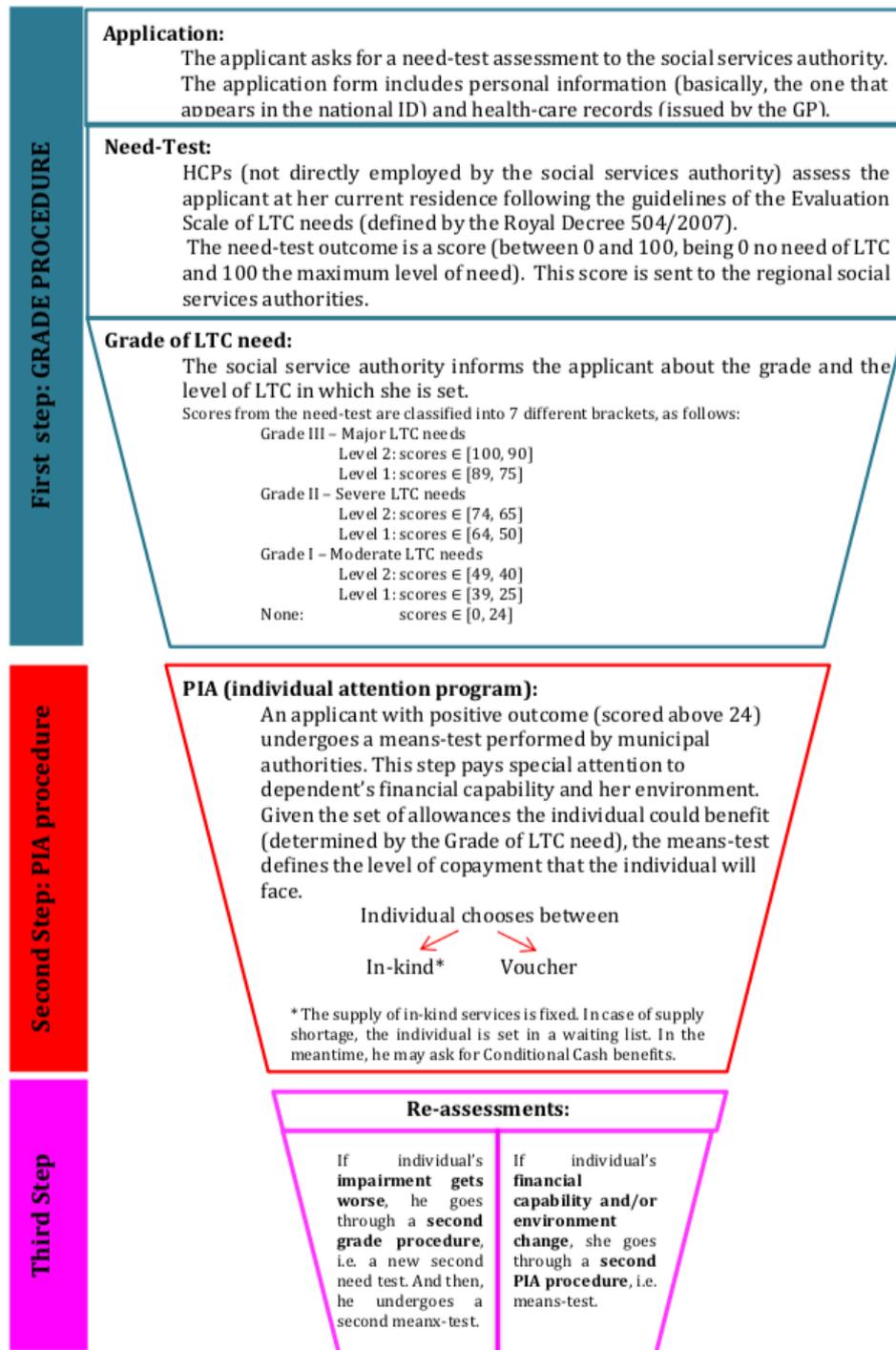
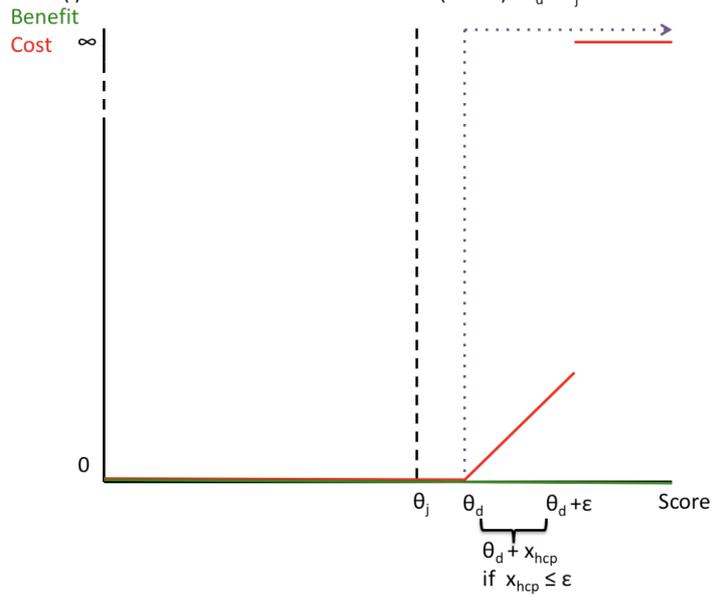
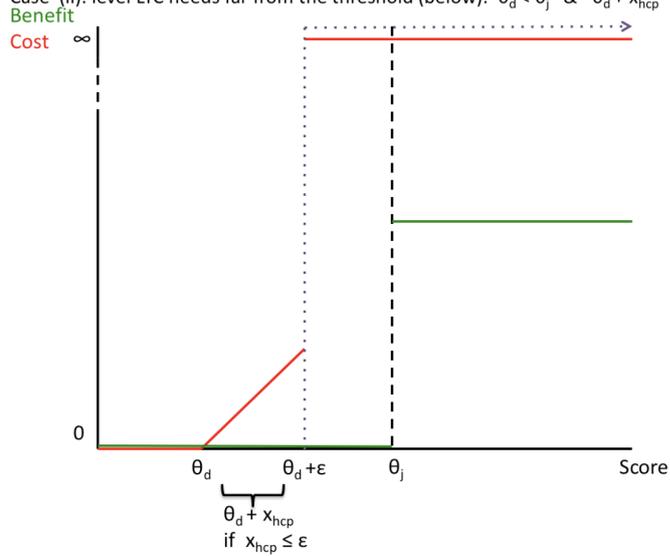


Figure A.2: **Prosocial HCP**

Case (i): level LTC needs far from the threshold (above): $\theta_d > \theta_j$



Case (ii): level LTC needs far from the threshold (below): $\theta_d < \theta_j$ & $\theta_d + x_{hcp} < \theta_j$ & $0 \leq x_p \leq \epsilon$



Case (iii): level LTC needs close to the threshold: $\theta_d < \theta_j$ & $\theta_d + x_{hcp} \geq \theta_j$ & $0 \leq x_{hcp} \leq \epsilon$

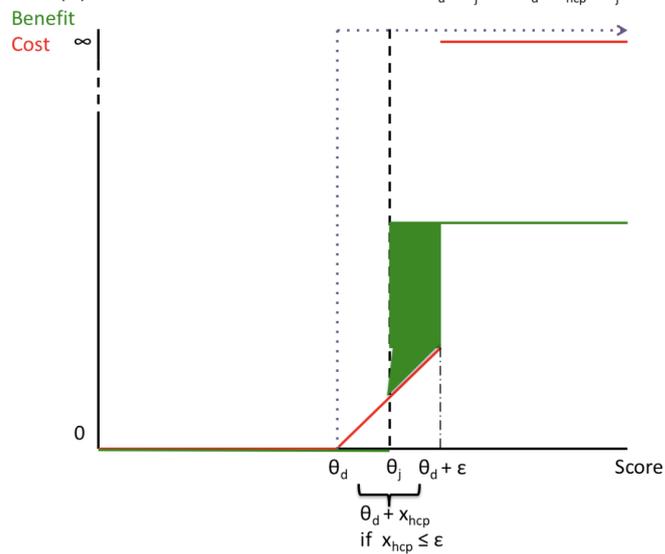


Figure A.3: **Threshold Shift: the unintended consequence of a discrete scheme of benefits with prosocial HCP**

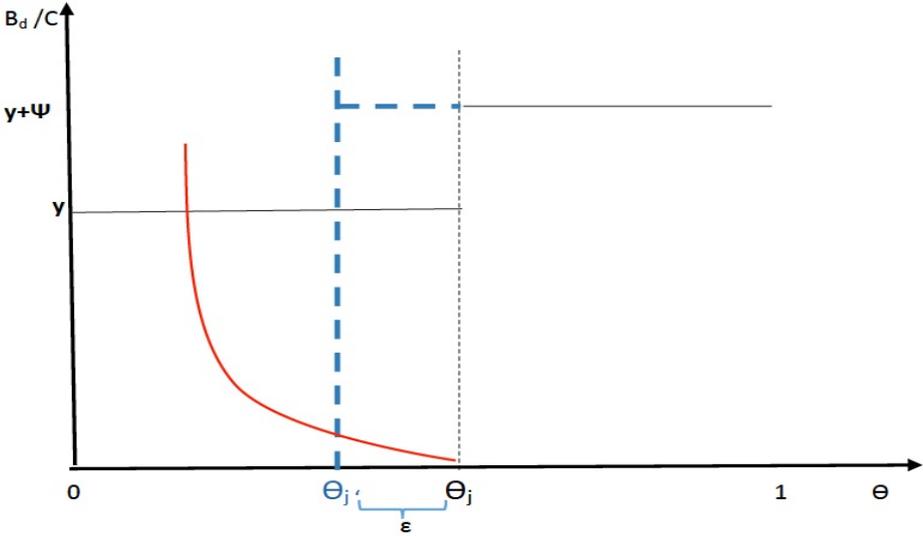


Table A.1: **Monthly Monetary Value of LTC Benefits**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Informal Caregiver (IC)	Nursing Homes (NH)		Day Care Centre (DCC)		Home Care (HC)	TeleAssistance (TA)
<i>type of benefit</i>	<i>cash transfer</i>	<i>voucher</i>	<i>in-kind</i>	<i>voucher</i>	<i>in-kind</i>	<i>in-kind</i>	<i>voucher</i>
Grade III	431	831	1870-c	409	853-c	537	
Grade II	303	494	1595-c	247	730-c	307	20-c
Grade I	168			171	597-c	211	

Notes: All amounts are in euros 2012. For the benefits in *voucher* or *cash transfers* the reported amount is the average, as the amount depends on beneficiary's financial capability. For *in-kind* benefits, the monthly value is defined as the public cost/price of the service minus the copayment (C), which depends on beneficiary's financial capability.

Table A.2: *Adjusted claimants and estimated parameters, by years and threshold*

(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>From none to Grade I-I: 23-27</i>						
$\Delta 23$	$\Delta 24$	$\Delta 25$	$\Delta 26$	$\Delta 27$	λ	θ
2008	255	285	-222	-134	-184	0.067 479.40
2009	474	559	-485	-293	-255	0.088 816.00
2010	487	549	-466	-344	-226	0.065 819.00
2011	251	315	-246	-182	-138	0.128 500.80
<i>From Grade I-I to Grade I-II: 38-42</i>						
$\Delta 38$	$\Delta 39$	$\Delta 40$	$\Delta 41$	$\Delta 42$	λ	θ
2008	192	215	-159	-128	-120	0.072 624.61
2009	361	410	-370	-225	-176	0.064 950.40
2010	379	646	-515	-336	-174	0.067 973.67
2011	224	294	-293	-139	-86	0.123 567.00
<i>From Grade I-II to Grade II-I: 48-52</i>						
$\Delta 48$	$\Delta 49$	$\Delta 50$	$\Delta 51$	$\Delta 52$	λ	θ
2008	527	604	-1043	-119	31	0.039 1021.60
2009	707	846	-1075	-381	-97	0.067 1257.18
2010	433	916	-808	-367	-174	0.099 1086.00
2011	235	290	-278	-179	-68	0.100 556.20
<i>From Grade II-I to Grade II-II: 63-67</i>						
$\Delta 63$	$\Delta 64$	$\Delta 65$	$\Delta 66$	$\Delta 67$	λ	θ
2008	620	701	-702	-411	-208	0.058 1223.60
2009	500	574	-645	-326	-103	0.058 1107.80
2010	151	535	-444	-237	-5	0.112 837.80
2011	86	119	-137	-64	-4	0.122 374.00
<i>From Grade II-II to Grade III-I: 73-77</i>						
$\Delta 73$	$\Delta 74$	$\Delta 75$	$\Delta 76$	$\Delta 77$	λ	θ
2008	476	530	-540	-297	-169	0.050 1368.40
2009	245	321	-428	-132	-6	0.090 963.20
2010	35	322	-295	-93	31	0.115 631.20
2011	18	39	-72	-1	16	0.180 283.40
<i>From Grade III-I to Grade III-II: 88-92</i>						
$\Delta 88$	$\Delta 89$	$\Delta 90$	$\Delta 91$	$\Delta 92$	λ	θ
2008	532	621	-823	-161	-169	0.055 1828.80
2009	182	220	-430	-33	61	0.039 869.00
2010	19	177	-201	-18	23	0.094 429.80
2011	6	12	-45	-4	31	0.058 183.00